

**Market Participation and Embedded Critique in
John Rawls' Theory of Justice**

Craig Duckworth

University College London, Ph.D.

UMI Number: U592791

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U592791

Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract: John Rawls has been criticised for wrongly specifying the relationship between persons' ongoing emotional and intellectual commitments and their capacity to reflect on and revise those commitments. While there are, arguably successful, responses to this critique, the difficulty shows up, I argue, in connection with his representation of the problem of stability. Stability, in an ongoing Rawlsian society, if it is to be realistic, requires an accommodation of competing, personal concerns, concerns grounded in "comprehensive" moral doctrines. However, when appeal to Rawls' principles of justice is required in order to settle disputes, in an ongoing Rawlsian society, the disputants must adopt a neutral standpoint that mutes the practical salience of their personal concerns. Those concerns, then, will not be engaged in deliberation. This has the implication that a person must respond to a judgement that goes against her by rescinding what she sees, *prima facie*, as a legitimate concern. This represents the problem of stability as one of how to extinguish unsupported practical considerations. The problem ought to be represented, rather, as one of how to accommodate competing concerns while keeping their practical salience, for the individuals whose concerns they are, intact. Ways to attempt to remedy this can be drawn on from outside of Rawlsian theory. In this thesis, however, I attempt to show how one can address the issue from within a Rawlsian account. Centrally, I note that Rawlsian citizens are, by assumption, market participants, at least insofar as they are involved in a system of discretionary exchange, and that, in order to be adequately specified as market participants, they must possess certain characteristics. An exploration of these characteristics offers, or so I argue, a conception of the relationship between ongoing personal commitments and the capacity for radical critique that promises to address the deficiency in Rawls' representation of the problem of social stability.

Contents

Introduction.....	5
 Chapter 1: A Problem with the Problem of Stability in Rawls.....	14
§1 A Response to the Neo-Hegelian Critique of Rawls.....	16
§2 Stability, Quasi-Stability, and Characteristic C.....	22
§3 Stating the Problem.....	33
§4 Stability and the “Oversocialised” Conception of the Person in Rawls.....	41
§5 The Artificial Defence.....	51
§5.1 Sandel and the Artificial Conception of the Person in Rawls.....	53
§5.2 Stability in a "Realistic Utopia".....	60
§6 Conclusion.....	62
 Chapter 2: The Embeddedness of Market-Oriented Choice: A Response to Cohen on Rawls.....	63
§1 Cohen's Thesis.....	67
§2. Questioning Williams’ Ethos-Based Reading of Cohen.....	70
§3. Pogge’s Outcome-Oriented Reading and a Reasons-Based Alternative.....	77
§4. Rejecting the Reasons-Based Version of Cohen’s View.....	88
§4.1 First aspect of the argument: The neutral basis of institutional principles.....	88
§4.2 Second Aspect of the Argument : The embeddedness of market-based decision.....	98
§5. A "Delimitative" Response.....	107
§6. Conclusion.....	111
 Chapter 3: Market Participation and Embedded Critique in Rawls.....	115
§1 The Market in Rawls.....	116
§2. Limited and Unlimited Reason-Based Autonomy in the Context of Market Exchange.....	123
§2.1 Reasons, Evaluative Criteria, and Preference.....	123
§2.2 Reason-Based Discretion in Dworkin's Resource Egalitarianism.....	125
§2.3 Reason-Based Discretion in Markets and Hierarchies.....	130

§3. The Argument Adopting a “Metapreference” Approach.....	139
§4. Implications for Cohen’s Argument?.....	141
§5. Conclusion.....	143
Addenda: I.....	145
II.....	147
Chapter 4: McDowell, Rawls, and the Market.....	154
§1. Embedded Critique as a Via Media Between Rawls and the Communitarians.....	158
§2. Williams and Taylor.....	165
§2.1 Williams.....	165
§2.2 Taylor.....	169
§3. McDowell’s Moral Philosophy.....	177
§3.1 The Epistemological Basis of McDowell’s Moral Philosophy.....	179
§3.2 McDowell’s Ethics.....	190
§4. A McDowellian Reading of Rawls.....	197
§5. Beyond the Market Context.....	207
§5.1 Metapreference and the Basic/Non-Basic Distinction.....	209
§5.2 McDowell and the Market Agent.....	212
§5.3 Beyond the Market Context.....	216
§6. McDowell, the Market, and Stability.....	218
§7. Conclusion.....	221
Conclusion.....	223
Bibliography.....	227

¹ Rawls, 1999/1971, §4.

² Rawls, 1999/1971, §24.

Introduction

John Rawls' *A Theory of Justice* (Rawls, 1999/1971, henceforth, *Theory*) presents a contractarian justification for liberal egalitarian institutions. A just society is, on this account, one whose legal, social, and economic policies are informed by principles that would be chosen, by its citizens, under conditions that are fair. The fairness of the hypothetical choice scenario (what Rawls calls the "original position"¹) is secured by denying information to the contractees that would tend to induce them to argue for principles that reflect personal bias. This "veil of ignorance"² is designed to model impartial reasoning, then, by removing information about persons' talents, social status, wealth, and the substance of their personal moral, philosophical, and religious beliefs. Social stability is of central importance in *Theory*, being one of a number of criteria with which to select principles of justice. It takes on even greater importance in Rawls' work when he comes to realise the incompatibility of the way he argues for the principles of justice and a presumption of radical disagreement across different cultural and ideologically differentiated social groups, about what constitutes the good.

Rawls' contractualist approach, as he presents it in *Theory* requires that principles for a just society be supported for reasons all can share as part of their conceptions of the good. This presumes a common evaluative basis for principles of justice that any group, given the radical nature of their disagreement with other

¹ Rawls, 1999/1971, §4.

² Rawls, 1999/1971, §24.

groups in the same society, ought to be able to reject, without thereby disqualifying themselves as Rawlsian citizens. Realising this, Rawls reworked his accommodation of stability and the fact of radical disagreement, the “fact of pluralism”³, culminating in *Political Liberalism* (1996/1993)⁴. In essence, whereas in *Theory* support for public institutions was to be understood as emanating from persons’ conceptions of the good, in *Political Liberalism* support for public institutions is made possible through the capacity people have for a mode of thinking that takes the political as its special domain, leaving conceptions of the good to inform evaluative thinking in personal contexts, and group-specific discussions. A further distinction between Rawls’ thesis in *Theory* and that of *Political Liberalism* is his emphasis, in the latter, on the artificiality of the conception of the person employed there. In this Rawls was responding to a style of criticism that objected to the way in which he characterised the moral point of view. According to this type of critic Rawls fails to embed the moral reasoning he requires for evaluation of existing institutions in the ongoing moral context of the people whose institutions they are. Stripped of the information associated with their own conceptions of the good people will derive normative principles that are arbitrary from the personal point of view, having no foothold in their everyday way of seeing things. There are a number of ways in which this type of criticism can be made but most useful, for my purposes here, is the following. The principles of justice Rawls derives are intended, in the first instance, for what he

³ That is, importantly, the fact of “reasonable” pluralism. A just, liberal society need not accommodate all conceptions of the good, but can reject those that, for example, threaten its viability (see Rawls, 1996/1993, 36).

calls a “well-ordered society”⁴, a society characterised by ideal conditions, concerning, for example, the extent of agents’ compliance. While this is an idealisation it is Rawls’ intention that the well-ordered society not be a distortion of reality, but be sufficiently realistic to present the bases for a realisable liberal egalitarian polity. He fails in this objective because in order for people to moderate their everyday conduct so as to act in accordance with normative principles of justice, such as those Rawls recommends, they must recognise the practical significance of those principles from within their everyday point of view. Because Rawls requires an impersonal evaluative stance that neutralises the practical role of personal beliefs, affections, and convictions, in the justification for the principles of justice, agents must bypass those factors in order to appropriately modify their behaviour. This misconstrues ethical reasoning as it fails to recognise the embeddedness of moral evaluation, and so fails to adequately model the ethical thinking, and, therefore, the moral possibilities, of real world agents. Rawls’ response to this critique is to point up a distinction between the everyday, evaluative stance of the agents in the well-ordered society, an evaluative stance that is informed by what he calls agents’ “comprehensive doctrines”, and the artificiality of the neutral stance agents adopt when evaluating their society’s institutions. The normative stance of real world agents is informed by their comprehensive doctrines, being constituted by their philosophical, moral, and religious convictions, and this is reflected in the conception of the evaluative thinking of well-ordered citizens in their everyday lives. The impersonal perspective they are presumed to adopt in an assessment of their society’s basic

⁴ Relevant references for this notion in Rawls can be found in chapter one of this essay.

institutions is not intended to reflect real world evaluative capacities but to function merely as a theoretical device intended to portray a form of moral thinking that does not presuppose a comprehensive doctrine. While we (real world agents) are capable of moral reasoning that does not call into play the convictions that underpin our comprehensive doctrines, representing this requires an artificial division of modes of evaluative thinking that, Rawls suggests, those who criticise him on this score mistake for an attempt to describe the actual evaluation of real world people. Rawls offers here an effective rebuttal. The realism of the well-ordered society need not be seen to be challenged by the use of a bipartite account of reflective thinking, once we recognise that the impersonality of the behind-the-veil point of view is merely artificial. Rawls' defence will not work, however, or so I want to argue, once we recognise the implications of the bipartite conception for his representation of the problem of social stability.

A stable liberal democratic society, in which stability is associated with non-violent, incremental institutional change, is marked by an ever-present tension between compliance and dissent. Non-violent critique requires that dissenters comply with existing institutions even while the latter are under review, and, where dissenters' claims are unsuccessful, that they, albeit reluctantly, comply with what prevailing institutions require. The problem of social stability, then, for the type of society that is Rawls' target, ought to be understood as accommodating an ongoing tension between what persons agree to do, with regard to institutional requirements, and what they would do were institutions structured in the way they desire. Dissent must be a live possibility for those who reluctantly comply with existing institutions. Rawls' conception of reflective

criticism will not allow for the ongoing possibility of a manifestation of dissent among those who reluctantly agree to abide by institutional requirements. Those who comply with existing rules do so by extinguishing, or obliterating the personal concerns that motivate their dissent, so as to remove the practical significance they would otherwise possess, in their everyday lives. This is because, in order to adjudicate between competing claims in the well-ordered society, the society regulated by Rawlsian principles of justice, it will be necessary, at least for certain hard cases, for agents to attempt to establish the standing of their claims by imaginatively adopting the stance of contractee in the original position. This requires that the agent adopt Rawls' bipartite conception of reflective criticism as a model for their own thinking. They then understand themselves to be evaluating the basis for their claims from a behind-the-veil position that puts out of mind the bases of the motive force of the personal concern that motivates their claim. This has them adopt a neutral stance to their personal concern that mutes its practical significance, this being unavailable from behind the veil. Stripped of their comprehensive doctrine in the neutral perspective of reflective evaluation, agents have no means to feed the results of such evaluation back into their everyday reasoning. They cannot moderate the motive power of their personal concern from within the point of view that is informed by their comprehensive doctrine because the mode of reasoning they have used to arrive at a normative assessment of their claim is alien to that point of view. In order to override personal concern, where that is required, a person must obliterate it from outside of his or her comprehensive doctrine, rather than moderate its motive power from within the doctrinal standpoint. Thus the

artificiality of Rawls' bipartite conception of reflective criticism no longer seems innocuous, once we recognise its implications for the understanding, in Rawls, of the problem of stability. The problem is understood to be one of how to root out, or eradicate, dissent rather than the problem of how to motivate agents to moderate the motive purchase of their personal concerns, in their everyday lives.

In this essay I do not attempt to solve the problem of social stability. Indeed the Rawlsian solution may be correct. What I attempt to do is to provide a conception of critical evaluation that allows for a representation of the problem of social stability that recognises the ongoing tension between compliance and dissent. I attempt to provide such a conception by drawing on, what I argue are, the characteristic marks of market participation. Everyday, personal, market-based choice is informed by comprehensive doctrine. This must be the case in a Rawlsian, well-ordered society, if Rawls is to preserve the distinction between the doctrinal basis of everyday choice and the neutral perspective associated with institutional evaluation. Dissent in the marketplace, associated with, for example, policy intended to root out unwarranted discrimination in employee selection, will involve a comparison, at the personal level, of competing considerations. The dissenter will, on one hand, be faced with a requirement that he or she choose in accordance with the institutional rule designed to effect Government policy, and will, on the other hand, possess personal criteria she would apply in the absence of the institutional rule. In order to arrive at a settled opinion as to which of these two, alternative evaluative criteria to apply, the market participant must bring the criteria into view. This perspective is one from which the presuppositions supporting the two types of criteria must not be presupposed. Were they to be

presupposed we would fail to distinguish between the neutral and doctrinal points of view that support the competing types of evaluation. The way is opened, then, to defend a conception of *presuppositionless criticality* that dissolves the divide between neutral and doctrinally committed points of view. This conception allows us to embed criticality because the neutrality required for impartial evaluation is possible without requiring the agent to stand outside of his or her comprehensive doctrine. Being so positioned, criticality need not entail the neutralisation of personal concern, leaving the latter's motivational pull open to moderation. This conception of embedded criticality does not solve the problem of how to achieve social stability, but it does offer a conceptualisation of the problem that can accommodate the ongoing tension between compliance and dissent that ought, I have suggested, to be thought to exist in Rawls' well-ordered society.

Market-oriented choice, though a central aspect of everyday life in a modern economy, covers only a limited domain of personal choice. It leaves out choice, and, more broadly, conduct, in family and other social contexts, and so the picture of embedded critique I have outlined here might be thought, for that reason, to be of limited interest to Rawlsian criticism. The conception of embedded critique associated with market participation has, however, implications for how we view the critical faculties of the Rawlsian citizen as such, and not only in his or her role as market participant. First, it might be thought that a presuppositionless conception of criticality, such as this, cannot to be associated with groups whose market conduct is exclusively informed by their group-specific norms. Such groups are market participants without being at all responsive to institutional requirements, it might be said. However, such groups are ruled out of

the Rawlsian well-ordered society because they lack the sensitivity to constraints on personal choice support for which comes from outside of their own norms. Moreover, restrictions on conduct intended to harmonise conceptions of the good will inform institutional design, and so being sensitive to institutional requirements, and not only their own group-based norms, will be required of all agents in a Rawlsian society. This being so, the market-oriented choice of all persons in a Rawlsian society must be informed not only by personal doctrine but also by what is prescribed (and proscribed) at an institutional level. Having established, then, that all (Rawlsian) agents possess the characteristic mark of market participation, a degree of criticality that transcends the division between neutral and non-neutral thinking, as that division is specified in Rawls, it is possible to view this as a mark of the critical capacities of the Rawlsian agent as such. It would be implausible to associate with the Rawlsian agent qua family member, for example, the bipartite conception of reflective critique many think constitutes a weakness in Rawls' thesis, while being convinced by the conception that, I am suggesting, ought to be associated with the agent qua market actor. It is possible, then, to argue for a more adequate picture of the place of dissent, in the conception of the problem of stability, by focusing first on the market context. The thesis can then be extended to the Rawlsian society at large.

In chapter one I lay out the problem with Rawls' representation of social stability, as I see it, relating it to a familiar style of Rawlsian criticism associated with a "neo-Hegelian" disapproval of the way he conceives reflective evaluation. Here the problem is framed in terms of a distinction between "oversocialised" and "undersocialised" conceptions of human agency. In chapter two I emphasise the

contrast between institutional constraints and personal doctrine as they influence choice in the market context. I do this via a response to G. A. Cohen's recent critique of Rawls. According to Cohen, Rawls cannot, as he would wish, restrict his principles of justice to background institutions, but must, necessarily, extend them to personal contexts, such as the market. I argue that Cohen fails to attach adequate weight to the role of comprehensive doctrine in everyday, market-based choice. Chapter three is an attempt to resolve the problem identified in chapter one. I attempt to show how a sound conception of embedded critique flows from a defensible understanding of certain characteristics of market participation. The central argument draws on the notion of "metapreference". In chapter four the work of John McDowell is brought into play. This is in order to support the view that it is appropriate to view the critical capacities of the market participant, and, more generally, the person, in a way that dissolves the divide between neutral and non-neutral points of view. I suggest ways, drawing on McDowell, in which the argument of chapter three might be extended to non-market contexts. A final section contains concluding remarks.

Chapter 1

A Problem with the Problem of Stability in Rawls

Introduction

John Rawls has been criticised for wrongly specifying the relationship between persons' ongoing emotional and intellectual commitments and their capacity to reflect on and revise those commitments. In this chapter I consider the way in which this (neo-Hegelian) style of critique shows up in Rawls' representation of the problem of social stability. I argue that, while there are, arguably, successful responses to the neo-Hegelian reproach, they do not succeed in the context of Rawls' representation of social stability. To put some preliminary flesh on the remarks I want to make, I note that the problem of stability ought to be represented as one of how to accommodate competing concerns and interests, on the understanding that the practical bearing of those concerns on personal conduct is a live, or ongoing, issue for the people whose concerns they are. A stable (or, at least, a stable *democratic*) society is one in which the possibility of dissent is an ever-present phenomenon. While individuals whose interests run contrary to prevailing institutions, and the restrictions on personal conduct they entail, will, if they are not actively subversive, restrain themselves, this does not mean that they will, nor that they are prepared to, radically overhaul their personal convictions, so as to remove their practical force. They will, rather, exercise personal restraint

in the respect they show for the existing institutional set-up, even though they would rather see alternative institutions prevail. Rawls, or so I will argue, fails to represent this tension, the tension between persons' commitment to existing policies, rules and procedures, and their wish that their society's institutions were constituted otherwise. While stability ought to be represented as involving ongoing temperance on the part of those who reluctantly accept existing institutional constraints, the way Rawls represents the problem has dissenters exorcise their personal commitments, commitments that do not sit well with existing institutions, and has them extirpate their normative significance.

In the sections that follow I state first (in section one) a response to the neo-Hegelian charge that Rawls' conception of reflective critique is faulty. In section two I note the role and character of stability in Rawls' theory, and argue that his conception of, what he calls, "quasi-stability" entails a particular conception of reflective critique: one that combines a capacity for critical reflection on prevailing institutions with ongoing obedience to their practical requirements. While this conception is represented differently in his earlier and later work - is improved upon, Rawls would say, in his later writing - neither successfully captures what is necessary for a realistic representation of social stability. As is argued in section three, the type of impersonal reasoning a Rawlsian person must engage in, in order to evaluate and defend the strength of his or her public claims, mutes the practical significance of the concerns that motivate her case. So subdued, the normative and practical weight of those personal concerns is lost in deliberation. So, where a person accepts, reluctantly, that her claim ought not, all things considered, to result in institutional adjustment, and so ought not to inform

her personal conduct, she will be unable to moderate (from within her everyday standpoint) the normative pull her concerns have for her in her everyday life. In order to suppress the sources of dissent the person will have to root them out, so as to be a person for whom such concerns simply do not figure in everyday life. Section four explores the roots of the problem, while in section five I address the point that Rawls might be immune to criticism in this regard because of his explicit intention to offer an artificial account of personal reflection, one that is not intended to mirror real world evaluative capacities. My argument, here, is that his misconstrual of the problem of stability runs counter to his "realistic utopianism". Section six concludes.

§1. A Response to the Neo-Hegelian Critique of Rawls

John Rawls' theory of justice, as stated, principally, in *A Theory of Justice* (1999/1971), *Political Liberalism* (1996/1993), and *Justice as Fairness: A Restatement* (2001), has come under attack from a number of, what have been called, neo-Hegelian critics. Prominent among these are Michael Sandel (1996, 1998/1982), Alisdair MacIntyre (1987/1981, 1984, 1988, 1999), Bernard Williams (1981a, 1985), Michael Walzer (1983, 1994, 2004), and Charles Taylor (1979, 1989).⁵ One strand in the neo-Hegelian critique of liberal political theory

⁵ Discussions can be found in, for example, Sen (1999b), Buchanan (1989), O'Neill (1988), and Kymlicka (1991). Feminist critics of liberal theory present similar objections. See, for example, Heckman (1992, 1995, Ch. 2). I choose the term neo-Hegelian rather than the more common "communitarian" in order to separate what it has to say about the liberal conception of the person from the associated communitarian political agenda. This allows me, *inter alia*, to include Bernard Williams in the list. As support for the view that there is a distinction to be had between the communitarian thesis and its political

draws attention to the role and character of personal autonomy in liberal thinking. A staple liberal view claims that personal and social interests are best served where individuals are allowed personal space to explore their ideas, inclinations, and talents. Thus the self is emphasised as a source of innovation and social renewal over the community and its collective energies. Supporting this emphasis on the individual is the liberal picture of the ideal agent: one capable of, and prepared to take up, a critical stance that doubts the validity of even the most deep-rooted of his society's norms and considered convictions.⁶ This, the neo-Hegelian argues, is a wrong-headed account of the critical distance a person is capable of when assessing her society's underlying principles, practices and routines. A person cannot simply shrug off her culture, its norms, its ready-made modes of evaluation, in order to think things through anew. The point is put vividly by Alisdair MacIntyre, who talks about the embeddedness of the Homeric character alongside the illegitimacy of supposing there to be a standpoint outside of the Homeric setting from which a self so situated could view his milieu:

implications, I note that Sandel's communitarianism supports his "republicanism" (Sandel, 1996) but, according to Larmore (1992/1987, 112), the one need not entail the other.

⁶ For other strands see, for example, Buchanan (1989) who identifies five currents running through neo-Hegelian views. The view I focus on in this essay is most akin to his point iv (p. 853); the view that "Liberalism presupposes a defective conception of the self, failing to recognize that the self is "embedded" in and partly constituted by communal commitments and values which are not objects of choice."

The self of the heroic age lacks..that characteristic which...some modern moral philosophers take to be an essential characteristic of human selfhood: the capacity to detach oneself from any particular standpoint or point of view, to step backwards, as it were, and view and judge that standpoint or point of view from the outside. In heroic society there is no 'outside' except that of the stranger. A man who tried to withdraw himself from his given position in heroic society would be engaged in the enterprise of trying to make himself disappear. (MacIntyre, 1987/1981, 126).

In similar vein, Michael Walzer argues that the social meaning of goods varies across distinct political communities, and that their social meaning forms the basis of a judgement about their appropriate distribution (Walzer, 1983). Hence an approach to justice that has goods *just as such* as its subject matter is bound to be mistaken by Walzer's lights. Indeed, a self specified in an acultural way so that all it can have in view are goods just as such would be similarly mistaken.⁷ And this is not solely a theoretical worry. A society that has such a (mistaken) self-understanding may have a severely morally impoverished public space⁸, or alienate its people because of a failure on the part of its government to endorse its citizen's beliefs.

A good way to characterise Rawls' project is as an attempt to combine radical moral, philosophical and religious plurality with a society-wide commitment to public institutions, sufficient to secure political stability over time.

⁷ (Cf. Walzer, 1994, x-xi: "There is a picture of the self, nothing so grand as a theory, that is consistent with "complex equality"" – a conception he explores in chapter five of that text).

⁸ Michael Sandel complains that the value-neutrality of public debate Rawls recommends creates a vacuum into which come the gossip columnist and the talk-show host (Sandel, 1998, 217).

And his recommended way of achieving this is to require agents to blind-side their personal normative commitments when assessing their society's institutions. In order to achieve this, Rawls attempts to provide a basis for the justification of principles of justice that can convince all of a society's members, irrespective of their standing, personal moral outlooks, or, what he calls, their "comprehensive doctrines".⁹ Such a basis is provided, in his thesis, through the construction of a fictional, contractual scenario, the "original position", in which the moral point of view is represented by placing carefully chosen limitations on the knowledge individuals have about their personal characteristics and social circumstance.¹⁰ In doing this, the neo-Hegelian charges, Rawls neutralises the emotional and motivational factors, the sensitivities and presuppositions that inform evaluative discussion within a living morality. An assessment of principles so derived, by the people whose lives they are intended to regulate, will require them to censor what, for them, are fundamental givens in moral discussion. Further, if, as the neo-Hegelian believes, such reflective neutrality engenders disenchantment then instability will be an inherent feature of a Rawlsian society. Stability requires civic virtue, fraternity and altruistic commitment, the argument goes, but these

⁹ Rawls, 1996/1993, 13: "A moral conception is...comprehensive when it includes conceptions of what is of value in human life, and ideals of personal character, as well as ideals of friendship and of familial and associational relationships, and much else that is to inform our conduct, and in the limit to our life as a whole. A conception is fully comprehensive if it covers all recognized values and virtues within one rather precisely articulated system...Many religious and philosophical doctrines aspire to be both general and comprehensive."

¹⁰ Eg. Rawls, 1973/1971, Chapter I, section 4.

depend on a "thick"¹¹ form of moral conviction that Rawls' account of personal reflection disallows. This, were it to be true, would substantially undermine his project. Further, and more fundamentally, if Rawls' specification of the possibilities of reflective thinking call for an impossible radical perspective then any conclusions that rely on it will require a different sort of support.

One response to this, offered by Rawls, is to emphasise the artificiality of the representation of the person in his theory of justice. The division in the person between the neutral stance she adopts in her assessment of public institutions, and the thick point of view she adopts in her daily life, models but does not attempt to depict how things actually are with real world evaluative psychologies. I return to this response in section five. A second way to respond is to redescribe the social world and persons' relationship to it in a way that responds to the neo-Hegelian critique but maintains the evaluative autonomy required for traditional, and Rawlsian, liberal positions. This approach acknowledges that group attachments, while not indissoluble, play a constitutive role in any correct characterisation of a person's normative outlook. According to this response we can accept the cultural embeddedness of persons, in a correctly specified liberal society, while arguing also that this ought not to lead us to give inadequate weight to the possibility each has of exiting his or her given social or cultural grouping. And associated with

¹¹ Walzer (1994, xi, n.1) uses the term "thick" to denote "a kind of argument that is...richly referential, culturally resonant, locked into a locally established symbolic system or network of meanings. He borrowed the term from "Clifford Geertz's defence of "thick description" in his much-cited *The Interpretation of Cultures* (New York: Basic Books, 1973)." In this essay I take "thick" to refer to the normativity-laden perspective a person adopts when her comprehensive vision is what is informing her thinking.

this possibility is the capacity for a critical distance that brings social and group-specific norms, and their philosophical basis, into view.¹² The neo-Hegelian ought, on this view, to accept this as a correct description of modern (liberal) societies unless she wants to disallow, implausibly, changes in normative stances, the possibility of movement among social groups, or reflective loyalty. And this degree of critical distance is all the liberal needs in order to derive principles for society as a whole, even given the fact of radical plurality.

It is commonly accepted that these responses are adequate to silence the neo-Hegelian critic. Rawls himself, while emphasising the artificiality of his conception, applauds Kymlicka's view that the liberal has anyway the resources to accommodate cultural embeddedness. And it seems right that we need not think of there being a radical gap, *in fact*, between personal and impersonal normative stances in order to theorise the scrutiny of shared institutions. It seems right, further, that modelling this form of reasoning will require a separation, somehow, of thick and thin commitments that need not be taken as depicting how things are with the evaluative reasoning of actual people.

¹² This type of response is found quite explicitly in Kymlicka (1991) and, can be borrowed also from Walzer (2004), even though, as is well known, Walzer is not in the habit of defending Rawls. In this recent work Walzer emphasises, among others things, the reluctance many people have to abandon their social roots even though an opportunity may have arisen to climb the social ladder. Conceptualising this requires that we characterise such a person as capable of reflective distance and emotional attachment, simultaneously. This picture can be imported, in principle, into Rawls' account without the rest of Walzer's methodological baggage. For a brief but illuminating account of the difference in method between Rawls and Walzer, see Carens (2000, 22-23).

However, when applied to the problem of stability in Rawls the neo-Hegelian critic is not so easily silenced, or so I want to suggest. There are two aspects to this critique and I prepare the way for them by looking, next, at the meaning of stability in Rawls. It will be seen that there is a relationship between the way he characterises stability and a conception of the reflective capacities of the Rawlsian agent. His theory incorporates this character to some extent, but not sufficiently, I will argue, to represent the problem of stability in a fully satisfying way.

§2. Stability, Quasi-Stability, and Characteristic C

Underlining the importance of the notion of social stability to Rawls' thesis is his early effort, in *A Theory of Justice* (henceforth, *Theory*) to define stability, and to express its significance:

Some measure of agreement in conceptions of justice is..not the only prerequisite for a viable human community. There are other fundamental social problems, in particular those of coordination, efficiency, and stability...(T)he scheme of social cooperation must be stable: it must be more or less complied with and its basic rules willingly acted upon; and when infractions occur, stabilizing forces should exist that prevent further violations and tend to restore the arrangement. (Rawls, 1999/1971, 5-6).

We cannot, in general, assess a conception of justice by its distributive role alone...We must take into account its wider connections...other things equal, one conception of justice is preferable to another when its broader consequences are more desirable. (Rawls, 1999/1971, 6).¹³

¹³ Cf. (Rawls, 1999/1971, 49, 119, 153-155; 1996, xix,).

Indeed, as is well known, Rawls did not come to think that he had down-played the issue of stability in *Theory*. Rather, he came to see that his account of the bases of social stability, in section III of that text, was inadequate once the implications of radical (though "reasonable") disagreements about social and human goods, among members of the same society, were fully recognised. In *Theory* Rawls sees the task of demonstrating the stability of a society regulated by the two principles of justice as one of showing the "congruence" (eg. Rawls, 1999/1971, 496) of justice and goodness. That is, he thought it was necessary to show that the just society, the one endorsed by "justice as fairness", would (or could) also be endorsed by its citizens from the perspective of their own conception of the good. Indeed, he says,

Best of all, a theory should present a description of an ideally just state of affairs...such that the aspiration to realize this state of affairs, and to maintain it in being, answers to our good and is continuous with our natural sentiments. (Rawls, 1999/1971, 417).

But this point of view he rejects as he comes to appreciate how "unrealistic"¹⁴ it is for societies marked by what he calls "the fact of reasonable pluralism"¹⁵, societies marked by a diversity of beliefs about the nature of the good. It cannot be presumed, that is, that all of such a society's members will endorse institutions as just from the perspective of their otherwise radically disparate conceptions of

¹⁴ Rawls, 1996, xvii, xviii.

¹⁵ Eg. Rawls, 1996, 24, n.27, 144.

the good.¹⁶ Stability in *Political Liberalism*, then, is founded not on a correspondence of conceptions of the good and the rational basis of just institutions, but is based rather on a capacity persons have to apply a separate mode of evaluation when assessing the justness of their society's basic institutions. This conception is presented perhaps most clearly in the following (Cf. Rawls, 1996, 12):

(J)ustice as fairness is to be understood at the first stage of its exposition as a free-standing view that expresses a political conception of justice...(T)he political conception is a module, an essential constituent part, that in different ways fits into and can be

¹⁶ "Conception of the good" is a rather vague notion. Rawls puts some flesh on it when he says that a conception of the good, "must not be understood narrowly but rather as including a conception of what is valuable in human life. Thus, a conception of the good normally consists of a more or less determinate scheme of final ends, that is, ends we want to realize for their own sake, as well as attachments to other persons and loyalties to various groups and associations...We also connect with such a conception a view of our relation to the world – religious, philosophical, and moral – by reference to which the value and significance of our ends and attachments are understood." (Rawls, 1996, 19-20). In what follows I often use Rawls' term "comprehensive doctrine" in place of conception of the good, though their relationship is confusing (see Rawls, 1996, 13). It seems that, a conception of the good should be identified with a "fully comprehensive doctrine", where the latter can be partial or full depending on the range of subjects to which it is applied. Also, a conception of the good takes for its subject matter all areas of political and personal life, including background institutions. A capacity on the part of citizens of a democratic society for the neutral thinking that is appropriate to the evaluation of public institutions that enables them to bracket their conceptions of the good for this domain. Rawls notes, in his discussion of public reason: "different procedures and methods are appropriate to different conceptions of themselves held by individuals...given the different conditions under which their reasoning is carried out, as well as the different constraints to which their reasoning is subject." Rawls, 1996, 220-221).

supported by various reasonable comprehensive doctrines that endure in the society regulated by it.

It is not suggested here that a justification for a commitment to just institutions should be viewed as grounded in the mode of evaluation associated with a person's religious, moral, and/or philosophical convictions. And that is, of course, a central point of departure from the point of view Rawls put forward in *Theory* (Cf. Rawls, 1996, 388, n.21).

There are a number of difficulties with Rawls' approach to stability in both earlier and later work. One concerns the reconciliation of his ideal-theoretic assumption of "strict compliance" (Rawls, 1999/1971, 8) with the possibility of non-compliance that underpins the need to secure stable institutions. An attempt at reconciliation here is instructive. While the problem of stability, as Rawls defines it (Rawls, 1999/1971, 12, cited above) is a problem for real world societies it seems as if Rawls' ideal-theoretic approach to its resolution precludes stability, so defined, as a problem. The principles of justice he derives are intended, in the first instance, to apply to the basic institutions of a "well-ordered" society (Rawls, 1999/1971, 4 (also, p.397)). A well-ordered society is one in which,

everyone accepts and knows that the others accept the same principles of justice, and..the basic social institutions generally satisfy and are generally known to satisfy these principles.

(A) well-ordered society is also regulated by its public conception of justice. This fact *implies* that its members have a strong and normally effective desire to act as the

principles of justice require. Since a well-ordered society endures over time, its conception is *presumably* stable. (Rawls, 1999/1971, 398. My emphases).

However, this seeming tension, between a presumption of stability and its use as a criterion for a successful account of justice, can be dissolved once we notice two important (though little advertised) qualifications of the notion of stability Rawls employs in his theory. First, he describes the type of social stability he intends as "quasi-stability":

A well-ordered society is quasi-stable with respect to the justice of its institutions and the sense of justice needed to maintain this condition. While a shift in social circumstances may render its institutions no longer just, in due course they are reformed as the situation requires, and justice is restored." (Rawls, 1999/1971, 400, n.3).

We are not, then to picture social stability as an equilibrium in the sense of an ongoing harmonisation of interests against a background of unchanging legal procedures and social and economic policies. Well-orderedness is compatible with gradual alteration to, or evolution in the structure of, background institutions. What is assumed away, in ideal-theoretic reasoning, is not a capacity to criticise what is considered, say, unjustifiably burdensome. What is assumed away is a preparedness on the part of society's members to consider the practical restrictions associated with existing institutions ineffective while they are subject to reflective critique. Quasi-stability, in the context of ideal theory, allows for internal criticism but requires non-radical shifts between institutional set-ups. This is compatible with a well-orderedness that presumes commitment to existing institutions (in terms of their practical import).

Second, Rawls does not desire stability *simpliciter*, but "As always...stability for the right reasons." (Rawls, 1996, xlii). Individuals in a well-ordered society must *volunteer* their assent to just institutions *because of* the rational appeal of their justification. Were this not the case, the society, "would not be liberal." (Rawls, 1996, 143):

(T)he problem of stability is not that of bringing others who reject a conception to share it, or to act in accordance with it, by workable sanctions...Rather justice as fairness is not reasonable in the first place unless in a suitable way it can win its support by addressing each citizen's reason.." (Rawls, 1996, 143).¹⁷

¹⁷ Rawls justifies his decision to approach things via "strict compliance theory", (Rawls, 1999/1971, 8) by claiming that it provides "the only basis for the systematic grasp of...more pressing problems," (ibid.) – those associated with, for example, civil disobedience. This suggests that we can safely distinguish between the possibility of critique (possible in a well-ordered society) and subversive conduct (precluded by well-orderedness). The contrasts here between what is and is not possible in a well-ordered society support a response to Wingenbach's complaints about Rawls' account of social stability (Wingenbach, 1999). He suggests, among other things, that Rawls' notion of stability fails to accommodate conflict over conceptions of justice (Wingenbach, 1999, 14), and this seems correct. Evaluation of existing institutions does not entail evaluation of the background conception that informs the principles of justice in the first stage of reflection on existing public rules. However, Rawls distinguishes early in *Theory* (p.5) between disagreement over conceptions of justice in actual societies and its presumed absence in the well-ordered society. It is not an implication of his theory, then, that he does not accommodate disagreement of this sort, it is an assumption. Wingenbach also says that Rawls' account, "fails to distinguish between justice and legal order" (Wingenbach, 1999, 14). While this appears to be true, given Rawls' definition of stability (Rawls 1999/1971, 12), once we make his notion of stability compatible with his ideal assumptions, an adequate distinction appears between legal bounds on behaviour and stability as ongoing, voluntary, and conditional, assent. In addition, the distinction Rawls sees between quasi-stability and stability *per se* appears in a more general context, in Ake (1974, 584-585): "(W)ritings on modernisation...conceptualize political instability

The voluntariness associated with quasi-stability, entails non-violent transition between institutional frameworks, and this conception of stability entails a particular characterisation of the relationship between a person's standing commitment to existing practices and her capacity for a critical response to them. Indeed, it is in the way this characterisation changes between *Theory* and *Political Liberalism* that a contrast between those works can be formulated. It also helps to identify a weakness in Rawls' conception of stability that transcends the issue, addressed in *Political Liberalism*, of the legitimate grounding of persons' commitment to just institutions.

Note, first, that there are different ways in which a person (in a well-ordered society) might criticise her society's existing institutions. Specific policies might be criticised because, while in intention they satisfy agreed upon principles of justice, they fail, in fact, to achieve their objective. A particular tax schedule, for example, might be thought a weak attempt to achieve equality according to the difference principle, though that that is its intention is not in dispute. A second possibility is that a policy might be thought to conflict with, or fail to adequately reflect, the background, regulative principles of justice themselves. A policy concerning the provision of faith-based education, that gives privileged funding

as, implicitly or explicitly, as violence, conflict, civil disorder, short duration of governments, lack of institutionalisation, and so forth...These conceptions do not fully reflect the fact that, properly speaking, only relations or a system of relations can be said to be stable...The application of the term (to a relation) always suggests the possibility of change in that relation, a certain possibility of centrifugence among the elements which constitute the system." Wingenbach appears to work with Rawls' basic definition

status to particular faiths might, for example, be argued to be inconsistent with the principle that a society ought to secure equality of opportunity. Third, in the manner of an exercise in reflective equilibrium, a person might want to evaluate the justice of her society's institutional backdrop *in toto*. Rawls (1996, 28) appears to deny that agents in a well-ordered society *can* engage in a process of reflective equilibrium. This is only open, he says, to "you and me" (i.e., us and Rawls) (ibid.), but I don't see how this can be sustained, as the justness of a well-ordered society's institutions are primarily a matter for the agents in that ideal world. This is not, however, of central importance for what follows. It is the second type of possibility that is instructive here: criticism of a specific, existing policy that questions whether it is justified by the reasoning that informs the principles of justice. It will, as Rawls says (1996/1993, 229) be difficult on occasion to determine whether actual policy is so justified:

(W)hether the aims of the principles covering social and economic inequalities are realized is...difficult to ascertain. These matters are nearly always open to wide differences of reasonable opinion; they rest on complicated inferences and intuitive judgements that require us to assess complex social and economic information about topics poorly understood.

This does not, however, preclude an attempt to settle disputes (in the well-ordered society, regulated by Rawlsian principles) by invoking the justification for the regulative principles of justice themselves, those that, it is agreed, ought to inform policy design. Now, a person who engages in such criticism, criticism of existing

(199/1971, 12) which is more akin to the definition Ake ascribes to "writings on modernisation".

economic and social policy, must, in order for her criticism to be compatible with quasi-stability, and voluntariness of commitment, have the following core characteristic, (C):

(C) She must be able to engage in counterfactual critique of existing institutions while respecting the practical requirements associated with those institutions as they stand.

To be clear, stability in the well-ordered society that Rawls describes does not require a once-for-all commitment to a particular set of basic institutions. While he assumes “strict compliance”, universal, ongoing obedience to the institutions that the Rawlsian principles of justice support, he recognises it may be difficult to achieve universal agreement on the way actual institutions are designed. Moreover, a flaw might be identified in existing institutions that, all agree, suggests a need for institutional redesign. So, while agents comply with institutions and the constraints on personal conduct they entail they, nonetheless, appreciate the contingency of those constraints, and recognise that the arguments in favour of existing institutions, and the constraints they entail, are questionable. Obedience to the constraints associated with existing institutions does not preclude a counterfactual stance that considers what society would be like were existing institutions otherwise constituted. Characteristic C is intended to capture what must be true of an agent if he or she is to be, at once, both non-subversive and reflective in her obedience.

The existing, personal, practical implications of institutions inform the individual's personal conduct as someone engaged in everyday life in the well-ordered, up-and-running, Rawlsian society. Critical reflection that takes into account the soundness of the justificatory ground for existing policy, on the other hand, requires a standpoint that transcends the everyday point of view. It must be the case, however, that while the transcendent (behind-the-veil) point of view is adopted the practical import of existing institutions is, nonetheless, in place, and in *Theory* this core characteristic is achieved by weaving together persons' "sense of justice" and their personal good. It is mainly in chapter eight of *Theory* that Rawls presents a sketch of the moral development of the person from infancy to adult maturity¹⁸, culminating (p.415) in an account of a mature "sense of justice". There he says that,

a sense of justice shows itself in at least two ways. First, it leads us to accept the just institutions that apply to us and from which we and our associates have benefited. We want to do our part in maintaining these arrangements...(...)..Secondly, a sense of justice gives rise to a willingness to work for...the setting up of just institutions, and for the reform of existing ones when justice requires it.

Furthermore, we are not to suppose that a person, who on reflection, sees the justness of the institutions that shape his action, takes on board the entailed personal requirements as external constraints on his behaviour. A person with a developed sense of justice embraces the requirements on personal conduct,

¹⁸ Distilled into three "psychological laws" at page 429.

entailed by institutional rules, as part of the fabric of his personal moral vision.

The adult citizen in a well-ordered society is one for whom,

Ethical norms are no longer (as in childhood) experienced merely as constraints, but are tied together into one coherent conception.” (Rawls, 1999/1971, 434).

Hence, a preparedness to work for the reform of existing institutions “when justice requires it” (ibid., p.415) is combined with an internalisation of existing constraints that embeds them in existing normative commitments. Being woven into the fabric of persons’ conceptions of the good, institutional constraints will be respected even while they are subject to critical review, while their justness is under suspicion. Rawls comes to see this way of harmonising compliance and critique as incompatible with a recognition of radical plurality. That is, given that the universal sense of justice emanates from disparate conceptions of the good, and Rawls comes to see that this pays insufficient respect to the fact of pluralism, he adjusts the way he incorporates core characteristic C in *Political Liberalism*. There a commitment to criticise institutions while respecting their practical authority derives from a capacity for, and preparedness to engage in, neutral thinking as regards basic institutions, from a capacity, that is, to engage in a mode of reasoning appropriate to political issues, that does not invoke “comprehensive” points of view (recall n.12, above).

It appears, then, that Rawls, were he to recognise characteristic C as a fair implication of quasi-stability, would see his conceptions of reflective critique, in both *Theory* and *Political Liberalism*, as accommodating it. In *Theory* existing institutions are wedded to existing conceptions of the good and so are supported

even while they are subject to criticism. In *Political Liberalism* a mode of appraisal appropriate to political discussion allows for ongoing commitment to existing institutions while they are under review. However, there is a problem here. Characteristic C, as stated above, does not capture all of the implications of the combination of quasi-stability and voluntariness of commitment for the conception of personal reflective critique. I try to state the nature of this difficulty in section three, before, in section four, exploring its roots.

§3. Stating the Problem

Consider an unhappy Rawlsian. He is a citizen of a well-ordered society, regulated by the two principles of justice, who will only make reasonable claims, regarding what he views as unfair burdens, or unfair distributive shares. He is unhappy because he thinks he perceives an injustice to him, and to his community. I will put some hypothetical flesh on why he is disgruntled in a moment. Note for now the following characteristics that a person so disgruntled must display, if he is to possess characteristic C. He is committed to existing institutions, only insofar as they reflect the principles of justice. While he is obedient, then, to institutional constraints, his commitment to them is inherently critical. Were it not then he would not be alive to occasions on which “justice requires” (Rawls, 1999/1971, 415) institutional reform. This critical acceptance of institutions is part of what makes a person’s commitment voluntary, and his institutions liberal institutions. Furthermore, acceptance of existing institutions will be *critical acceptance* because, while he accepts what they entail for his personal conduct, or, if he is public spirited, what they entail for society as a

whole, existing institutions run counter to, or only partially fulfil, his own aspirations. Quasi-stability and voluntariness entail a criticality, then, in the acceptance of existing institutions that places a tension between compliance and critique that Rawls, I want to argue, does not accommodate. While his theory, in both its earlier and later forms, appears well-placed to capture characteristic C, as I suggested towards the end of section three, above, it is not, I will argue, well placed to capture the mentioned tension.

Consider the following scenario. In a post-contract (up-and-running) Rawlsian society there is a minority cultural grouping (group 1) that has secured, through intelligent financial investment, say, greater resources with which to expand the number of opportunities it has to publicly celebrate its beliefs and traditions. Another minority group (group 2), with different beliefs and traditions, worries that the first group's raised profile may reduce social and economic opportunities for its (group 2's) members. The argument group that two puts forward in support of its claim is that an enhanced position for members of group 1, in the public imagination, a greater understanding of its traditions and rituals, will give it (group 1) an unfair advantage in the labour market, stifling group 2's efforts to advance itself. This will then, the argument might proceed, reduce group 2's ability to influence public debate because of its inability to support and garner general support for candidates to represent its interests.¹⁹ Both groups, we can assume, are small relative to other groups and a resource expansion for either would have to be very large for it to have a significant impact on the opportunities

¹⁹ Cf. Rawls, 1999/1971, 197-198: This would seem to be a claim that a Rawlsian government ought not to be deaf to. Resource inequalities, Rawls says, ought not to be such as to be the cause of inequality of political voice.

facing other cultural groupings in the post-contract society. So, significantly, group 2 is the only one making a claim for extra resources to combat what it sees as the threat from group 1. This is an important fact because, while the size of group 2, the number of its members, is relevant to the question why it and not other (larger) groups are concerned about group 1's economic progress, the reason why group 2 is making a claim, and not other groups, is that group 2 has what might be called a *heart-felt* concern. To share their concern, not in a detached, but in a heart-felt way, one would have to be a member of group 2²⁰. First, consider the way in which a typical member of group 2 must air his concern if he is to challenge the permissibility of group 2's economic advance.

Suppose group 2's claim finds no support in the existing constitution or the details of social and economic policy. Then the group will want to show that their claim would be accepted were the government's decisions informed by a correct understanding of the implications of the two principles of justice. They will want to show, that is, that a decision in their favour would follow from a general rule itself informed by the, more general, two principles. This, the second kind of critique mentioned on pages 23-24 above, is, it should be noted, consistent with the procedural nature of Rawls' theory of justice. The group are not claiming that a direct judgement can be made about their specific case by appeal to the two principles, only that a policy favourable to them, in its regulative character, can be

²⁰ According to the O.E.D. "heart-felt" can be defined as that which proceeds from "the innermost self". I use the term in something close to this strong sense in order to make vivid the connection between the claim made and the person's outlook. I refer, more generally, in the essay to persons' "personal" concerns. A concern does not have to be heart-felt, in the O.E.D sense, in order to be motivating.

so supported. But note the *reason* their claim will be successful here, if it is. If it is successful it will be because their claim is supported by a policy that in turn is supported by the more general two principles, that are in turn supported by behind-the-veil reasoning. Moreover, in order to argue that their claim is so supported the typical, or representative, member of group 2, who wants to make this argument, must adopt the perspective Rawls requires for impartial thinking: the behind-the-veil point of view that puts out of sight personal facts, such as, relevantly here, cultural identity. Along with cultural identity goes also, of course, the personal concerns whose basis is the cultural identity that is fully bracketed in the behind-the-veil point of view. There is a schism here. The reason for which group 2 is making the claim emanates from within their culture. It is not a reason available to all from an acultural standpoint, that is the point of calling their concern *heartfelt*, but operates as a reason only for those so culturally endowed as to see it as such. It is not a claim to additional resources just as such, and so a claim any might make irrespective of their cultural membership. Indeed, the group claiming disadvantage may accept that the resource inequality is due to good “option luck”, and so, arguably, fair by Rawlsian lights, yet still press their claim. So, on the one hand, in order to adequately ground their claim the group that feels the disadvantage must engage in argumentation that puts out of sight its cultural convictions in order to argue for public support for more resources. On the other hand, the reason for their claim is based not in considerations accessible to all from an acultural perspective. The considerations that give rise to the reason for their claim are intimately related to their cultural convictions.

At play here, *inter alia*, is the familiar neo-Hegelian criticism that the liberal reduction of the person to an individual as such puts a distance between her and the reasons she has for her more heartfelt claims. But more fundamental is the fact that the group cannot construct their argument, for a claim to public resources, in a way that has it *retain the practical salience* of the reasons that are internal to their way of seeing things. They must adopt a normatively neutral conception of themselves that mutes the motive power of their heartfelt concerns. A group can, of course, argue for a claim from behind the veil that refers to the importance to them of the relative social standing of their culture. But in order to do this, in the way that Rawls' original position device requires, they must speak of themselves not as the culturally loaded persons they are, with their cultural perspective intact, but as a cultural group as seen from the outside. They can, when they have adopted that point of view, refer to the reasons that initiate their claim, but they cannot give those reasons their practical, initiating role, that is they cannot keep in play the practical, motive status of the concerns, distinguished here by being called heartfelt. To give them this role they would have to be embraced from the standpoint of, in this case, group 2's cultural outlook.

And the point is a general one. Where the basic reasons for a dispute are culture-bound it is those reasons that are the motor for the dispute. But Rawlsian citizens are agents for whom such reasons are denied their initiating role, in an argument intended to settle disagreement among cultural groups. This does not put at an appropriate distance but entirely mutes the reasons that emanate from that source, so denying them their role as the practical ignition (so to speak) in a dispute. What is faulty here is not an implausible metaphysical split between a

situated self and its rational direction, nor the unreasonable expectation that people extinguish culture-based reasons in the formulation of their claims. The problem is that the requirement to begin with neutral considerations does not begin to engage personal reasons in disputation. It merely blind-sides them. It doesn't recognise and manage their force it merely extinguishes them, and in this way fails to give them their proper status, as the momentum in disagreement across cultural boundaries. What this entails is that when a person accepts a negative judgement, a judgement that goes against her, she has no way back in her thoughts to the practical reasons that were the basis of her claim. This is because the basis of those reasons, her heartfelt concerns, are put out of view in an impartial assessment that is constructed in the way Rawls constructs it in his model. Now, as an obedient, but disgruntled Rawlsian citizen, she must, having accepted the judgement, restrain herself so as to mute the practical purchase of her heartfelt concerns. But this she cannot do (cannot do, that is, in a way that addresses them with their motive status intact) because those concerns have not been engaged in deliberation. They must be extinguished or allowed free reign, but cannot, given that they have not been incorporated in evaluative thinking, be restrained as a matter of reluctant acceptance.²¹

Summarising this section, I note again that the problem of stability plays a pivotal role in Rawls' theory of justice. It can't be expected that groups with

²¹ I note that Sandel presents a similar thesis in one of his criticisms of Rawls: "The priority of the right over the good precludes members of different moral communities from bringing into play arguments *based on their peculiar convictions* in public discussion of "constitutional essentials and basic justice" (Sandel, 1998, 211. My emphasis).

differing doctrinal persuasions will commit to policies and procedures whose justification relies on normative presuppositions they do not share. For political stability there must be ongoing commitment to public institutions, and Rawls' objective is, in order to provide a model for actual polities, to show how in his "well-ordered" society such commitment can be sustained across doctrinally divided groups. It must be the case, then, that public procedures for dispute settlement, in a Rawlsian society, enable citizens to air their concerns and to work towards agreement. This is, in one sense, unproblematic. The constitutional phase of Rawls' theory of justice involves the implementation of legal principles and procedures, that, one might argue, must be adequate to dispute settlement, because they reflect the principles of justice Rawlsian citizens accept. This is not enough, however, to ensure that Rawls has successfully captured, or represented, the problem of stability in his conception of the well-ordered society. The importance disputants attach to their claims reflects their belief in the strength of those claims, even if it is accepted, also, that the claims they make may be faulty. That is, even where a critical stance is taken towards a conviction, and the typical Rawlsian citizen, being reasonable, can be expected to take such a critical view, the claim made will be *heartfelt*. It will stem, that is, from a world view, the commitments that constitute that view being the basis for the claim that is made. The basis of the claim must differ, then, from the basis of the Rawlsian person's support for the principles of justice. Were this not the case, the two would coincide, so undermining the gap between neutral and non-neutral stances so important in Rawls' theory. Yet, in order to press a claim so contentious that the principles of

justice must be brought into play - in order to work towards agreement - the discussants, in the well-ordered society, must adopt, forthwith, a stance that puts out of view, that negates the influence of, the heartfelt concerns that motivate the dispute. Part of the problem with this is that, detached from what they in fact care about, disputants clash not as citizens but as disengaged selves on whom personal concerns make no impression. But this is only part of the problem.

The problem of stability is a practical problem. It is the problem of how to deal in a single polity with the often competing and heartfelt claims of doctrinally divided, social groups; but the problem is represented, in the Rawlsian society, as one of how to arrive at agreement given a neutral standpoint all can share. On this basis, Rawls contends that we can overcome the problem of stability for a well-ordered society. However, he misdescribes the problem of stability by requiring agents to deal with disagreements by entering into a common, neutral standpoint that extinguishes rather than engages the heartfelt reasons of the parties to a dispute. If agreement is reached by means of such impersonal reasoning then what we have is a picture of a society from which sources of potential conflict have been expunged, rather than one in which those sources are tempered, in an ongoing accommodation of potentially disruptive interests. This undermines Rawls' intentions to present what he calls, in later work, a "realistic utopia" (Rawls, 2001, 4, 5-6, 11-12), a theory in which the properties of the ideal (well-ordered) society reflect what can be realistically achieved, given limitations on what can be expected, in terms of attachment and motivation, from real world agents. I return to this point in section five. In the following section I attempt to

further uncover the roots of the problem, and in so doing, to specify it more completely.

§4. Stability and the “Oversocialised” Conception of the Person

It will be helpful to attempt to uncover the root cause of the misrepresentation of the problem of stability in Rawls. A deeper understanding of the nature of the problem can be brought out by looking at the influential work of Mark Granovetter (1985). Granovetter's general argument relates to the sociologist's search for a more socially embedded, norm-laden alternative to the economist's isolated, rational actor. Applied to the marketplace Granovetter's concern is to emphasise the importance of ongoing social influences in market action. In arguing for his thesis Granovetter remarks on, what he calls, Rawls' "undersocialised" conception of the human agent (Granovetter, 1985, 483). In a way that is familiar from neo-Hegelian stances on Rawls' approach, Granovetter contends that Rawls pays too little attention to the role of socio-cultural norms in situating a person's moral outlook:

Actors do not behave or decide as atoms outside a social context...Their attempts at purposive action are instead embedded in concrete, ongoing systems of social relations. (Granovetter, 1985, 487).

Granovetter recognises the power of the liberal rejoinder to this critique, a rejoinder that points out the theoretical and real dangers of social determinism. If the defender of situated ethics is not careful to provide space for reflective critique, at the personal level, then personal freedom and autonomy are

threatened; and Granovetter shares this opposing concern. Drawing on Dennis Wrong (1961), Granovetter says he is also concerned about the converse consequences of an "oversocialised" rendition of the agent in terms of the limited space it might leave for social criticism at a personal level. What is interesting here is the way in which Granovetter thinks we ought to try to avoid the opposing pitfalls. I attempt next to bring this out. First, note that the "oversocialised" conception is, according to Granovetter,

A conception of people as overwhelmingly sensitive to the opinions of others and hence obedient to the dictates of consensually developed systems of norms and values, internalized through socialization, *so that obedience is not perceived as a burden.* (Granovetter, 1985, 483. My italics).

The italicised phrase in the quote helps to identify what makes Granovetter's thoughts distinctive. For him (p. 485) the central weakness in both undersocialised and oversocialised conceptions of Man is that neither captures the role of *ongoing* social relations and associated pressures in shaping personal conduct. Moreover, while some norms are associated with habitual behaviour, we ought to recognise that much social conduct, whilst it respects the pressures of social expectation, duty, and so on, involves an exercise of self-restraint on the part of the agent. The oversocialised conception, as Granovetter describes it, fails to make room for *the possibility of* self-direction that goes against expectation. Norms, moreover, are malleable things, and a responsiveness to, often subtle, changes in social norms needs to be built into the conception of the socially embedded person to reflect this. But the problem is not only that the person tends to be either undersocialised

or oversocialised, there is a relationship between the two (a relationship Granovetter doesn't, in fact, bring out explicitly).

Where a theory assumes, as does Rawls', according to Granovetter, an undersocialised conception of the person, for purposes of deriving normative requirements for the social context, it has difficulty representing agents, in their social context, as responsive to, and accepting of those requirements, rather than fully governed by them - "so that obedience is not perceived as a burden" (ibid.). Socialisation, on what Granovetter sees as a faulty account, is a matter of embracing norms that once embraced play an exclusive role in shaping choice. Once adopted they remove the need for an ongoing responsiveness on the part of the individual to her living, social context, because the once-for-all nature of the adoption of social norms has removed the possibility of acting in a way that is in tension with them. That is, they act on the individual as if from the outside, necessitating compliance, in a way that does not require any active social participation on the part of the individual agent:

Social influence is..an external force that, like the Deist's God, sets things in motion and has no further effects...Once we know in just what way an individual has been affected, ongoing social relations and structures are irrelevant. Social influences are all contained inside an individual's head, so, in actual decision situations, he or she can be atomized as any *Homo Economicus*... (Granovetter, 1985, 486).²²

²² This perspective has wide-ranging implications, according to Granovetter, for the analysis of economic activity. He argues, for example, that the "markets and hierarchies" literature associated prominently with Oliver Williamson relies on an artificial conception of the organisation as a place in which conformity is automatically forthcoming given hierarchical control. Against this one can point out that Williamson's theory allows for the

The point is made, perhaps more clearly, in Wrong:

What has happened is that internalization has imperceptibly been equated with "learning", or even with "habit formation" in the simplest sense. Thus when a norm is said to have been "internalized" by an individual, what is frequently meant is that he habitually both affirms it and conforms to it in his conduct. The whole stress on inner conflict, on the tension between powerful impulses and superego controls, the behavioural outcome of which cannot be prejudged, drops out of the picture. (Wrong, 1961, 187).

On this sort of account Rawls' agents are, in fact, at once both undersocialised and oversocialised. They are undersocialised because they are capable of imaginatively removing themselves from their social milieu and looking at things from the outside, as if everything, including the modes of evaluation that ordinarily inform their everyday conduct, is available for inspection. On the other hand, when the principles they accept are in play in their society they accept these in a way that has them override rather than engage their everyday commitments and beliefs. This makes them oversocialised in the lack of room it makes for the fact that new commitments ought to be characterised in such a way that they integrate with standing commitments, and, where appropriate, suppress, though do not extinguish, them. Moreover, the characterisation of the agents should be such as to keep in play the sensitivities

possibility of workers "shirking" in a hierarchy, though Williamson does seem to miss Granovetter's point when he refers (Williamson, 1998, 75) to "societal embeddedness" as a "background condition".

that enable them to commit but nonetheless take a critical stance towards their commitments. Without this there is no room for the tension between a commitment and the desires, dispositions, and reasoned intentions that push against it in everyday life.

For further clarification of the point, it will be helpful to consider what Granovetter (following Wrong) views as problematic in Thomas Hobbes' attempt to provide a model of social order (Hobbes, 1651/1985). Hobbes presents a contractualist account of a transition from a pre-social state, marked by the ever-present threat of violent disorder, to a peaceable, stable society, in which at the root of social order is compliance enforced by a central agent, the Sovereign. According to Granovetter, Hobbes' conception suffers the same weakness as Rawls' account. Rather than resolving the problem of the ever-present possibility of conflict that confronts Hobbes' pre-social agents, his proposed solution transports them immediately into a social scenario in which social conflict simply will not arise due to persons' unreflective obedience to normative requirements. In this way the problem of social order is not worked through, but merely removed, because as a norm-governed agent the potential in the person for following desires at odds with those consistent with existing social norms has been erased. So, when a person accepts the constraints imposed by sovereign authority in Hobbes' society he does so, on the Granovetter-Wrong thesis, on a model of internalisation of constraint that removes the possibility of him doing otherwise. And this, as with Rawls, is not an incidental problem for Hobbes. It is a result of the undersocialised conception of the person Hobbes works with. The only possible way in which internalisation can take place is through a radical transformation of

the way the individual thinks about his choices and what is permissibly open to him. The person that needed restraining is no longer in view. What is in view is a new person from whom the problematic characteristics – importantly, in the Hobbesian case, a readiness (supported by natural law) to resort to violence – have been removed as an ongoing concern. Thus the individual is transformed into an oversocialised personality incapable of having his inclinations vie with his impulse to obedience.²³

As an antidote for these perceived failings, Wrong (p.186) recommends Emile Durkheim's conception of the way in which social rules constrain individual behaviour, drawing attention to Talcott Parson's treatment of Durkheim

²³ The point is made also in connection with the conception of the person the (orthodox) economist adopts. Having a rational concern for what affects the self, the person can only be made sensitive to moral demands by appealing to her rational self-regard by showing her that moral boundaries are somehow better for her. This shows up in attempts to derive moral restraint in a game-theoretic, prisoner's dilemma framework - David Gauthier's "Morals by Agreement" (Gauthier (1987)) being one such, well known, attempt. The problem is, as Wrong sees it, not that moral restraint cannot be derived. The problem is that in showing moral restraint to be something that was anyway in the interest of the person they make the case for moral conduct completely compelling. For a person to see the possibility of immoral action as a live possibility for him makes him irrational and so disqualified from being an agent at all, on the definition of agency in play at the outset. Removing the possibility of immoral behaviour from the scope of a rational person's decision removes the problem the moral theory was intended to give an account of. Rather than retaining immoral conduct as a live possibility it removes it as a possibility for the rational person. Interestingly, Phillipa Foot's argument in "Morality as a System of Hypothetical Imperatives" (Foot, 1972) can be seen as an attempt to find a way out of the approach to moral theorising that sees moral demands as rationally compelling. The analogy she draws between morality and etiquette can be viewed as an attempt to introduce ongoing optionality (in a way that chimes with Granovetter's intentions) into a conception of moral agency.

in his *The Structure of Social Action* (Parsons, 1937, pp. 378-390). Durkheim came to think, Wrong tells us, of constraint arising from social rules not as an obstacle exerting pressure from outside of a person's psychology, but as permeating his personal reasoning, being, in fact, "constitutive" of the actor's ends. As Wrong notes,

Durkheim...recognised that the very existence of a social norm implies and even creates the possibility of its violation. (Wrong, 1961, 191).²⁴

In light of this discussion the problem with the representation of the problem of stability in Rawls can be viewed in the following way.²⁵ In order to enter into

²⁴ Wrong's primary concern, I should note, is with the conception of the person he felt was prevalent in sociology. A central task for the modern sociologist has been, he says, to provide an alternative to the motivational picture of the individual contained, for example, in utilitarianism, orthodox economics, and "the power-seeker of the Machiavellian tradition in political science." (Wrong, 1961, 190). It would be "ironical" he says (ibid.) if the sociologist's response, that Man has an overarching interest in group approval, should be merely "the creation of yet another reified abstraction in socialized man, the status-seeker of our contemporary sociologists." (ibid.).

²⁵ Vanberg (1994) identifies a central disagreement between economists and sociologists to be the sociologist's emphasis on rule-following and the economist's emphasis on choice in individual behaviour. He supports Wrong's reading of Parson's on Durkheim seeing the latter's intention to be a way of resolving the tension between a self-effacingly, norm-obedient conception of the person and one that sees her as continually opting to follow rules. For Parsons, Durkheim achieves this by integrating the constraints associated with rules and a person's extant interests: "the constraining factors actually enter into...concrete ends and values, in part determining them..." (Parsons 1960, 143,

dispute concerning hard cases, in the well-ordered society (where an issue's being a hard case entails that reference must be made to the principles of justice, and not just the procedures that interpret them), a person must adopt a stance that neutralises the motive power of the reasons he has for pressing the claim he makes. Suppose that his claim is unsuccessful because the principles of justice do not, on reflection, support his case. Then he will, if he is compliant, have to adjust his behaviour in order to abide by the judgement. But, because he has been required to adopt a neutral stance, and argue on that basis, there is no route from the judgement he accepts to a fine-tuning of his everyday way of seeing things. What he must do, if he is to abide by the judgement, is to transform himself, from the outside, in a way that does not engage, but bypasses, his comprehensive outlook.

The problem of stability, I have said, is a practical problem, a problem of how to accommodate, in an ongoing compromise, heartfelt claims, rather than of how to arrive at consensus given a shared neutrality. What is heartfelt is not tempered, or repressed, in an acceptance of an argument that draws on behind-the-veil reasoning. What is heartfelt is pushed aside, its vital force lost, in a radical transformation of the individual enlightened by what the principles of justice, he can now see, entail. To resolve this difficulty, neutral thinking, required of one who is reasonable enough to recognise the need to justify his claims to others, must be characterised in such a way that the practical salience, to him, of his concerns is somehow accommodated in deliberation. If those concerns can retain,

cited in Vanberg, 1994, 238). (Not all would agree that Parson's has captured Durkheim's meaning: the issue is discussed in Pope (1973, 1975), Cohen (1975), Parson's (1975)).

somehow, their vital status, then their motive force will, once deliberation is ended, be something that must, if required, be contained. What is blocked in such a conception is a presumption that personal concerns that have their basis in comprehensive doctrine have been rendered harmless in a person who accepts his claim is not justified. In chapters two and three I attempt to work up a response to this critique that draws on Rawls' presumption that his agents are market participants. It is important, however, to emphasise that the criticism I am making is not easily rebuffed by Rawls' stress, in his later work, on the artificiality of his conception of the person, and the associated model of reflective critique.²⁶

²⁶ In *Political Liberalism: Reply to Habermas* (1995) (reprinted as Lecture IX in *Political Liberalism*, 2nd edition) Rawls presents an account of his theory that suggests the divide, as I draw it, between comprehensive and neutral thinking does not well represent Rawls' view. In his reply to Habermas Rawls states that the principles of justice are *pro tanto*: "(S)ince political justification is *pro tanto*, it may be overridden by citizens' comprehensive doctrines once all values are tallied up." (Rawls, 1995, 143). In addition he says that, "We assume that each citizen affirms both a political conception and a comprehensive doctrine..In this case, the citizen accepts a political conception and fills out its justification by embedding in it some way to the citizen's comprehensive doctrine as either true or reasonable, depending on what that doctrine allows." (ibid.). Taken together these two statements imply that agents in a well-ordered society are to be understood to be capable of grasping at once both arguments with a neutral basis, and those that are associated with comprehensive doctrine. And, in order for agents to harmonise public reasons, those associated with arguments for regulative principles of justice, with comprehensive doctrine it must be the case that a standpoint is afforded them that allows a single evaluative perspective, one that dissolves the divide between comprehensive and neutral thinking. Such an all-embracing perspective would permit agents to raise personal concerns without annulling their motive force. On this reading, Rawls has an answer to both the neo-Hegelian charge that his conception of critical reflection relies on a metaphysically suspect level of rational detachment, and the objection that his agents are, as a result, oversocialised. It could be argued, then, that the

interpretation of Rawls I am working with is one that ought not to be maintained once we take into account the comments on his own theory Rawls makes in his reply to Habermas. I want to cast some doubt on that contention, while recognising that the point deserves more detailed discussion.

The pro tanto status of the principles of justice permits discussion of their rationale and allows that they are open to modification. However, when Rawls points this out, in his reply to Habermas, he notes also that the principles are to be discussed *from the point of view of agents* in a well-ordered society, “from the point of view of citizens in the culture of civil society” (Rawls, 1995, 140). And this is quite a dramatic change from Rawls’ earlier statement that the proper perspective for theoretical consideration of principles of justice is that of the theorist and his audience (Rawls, 1996/1993, 28). I think that this change is one that Rawls must, arguably, find hard to support.

Recall that according to Rawls the conception of the person employed in his theory is artificial. Are we to suppose, on this new account, that this is no longer the case and that the agents in the well-ordered society really do stand behind the veil of ignorance, from a higher-order perspective that embraces both that view and that associated with comprehensive doctrine? If so, the conception is no longer artificial and lets in the neo-Hegelian who wants to remark that this gets the metaphysics of normative reasoning wrong. Are we then to suppose that the model of reflective critique represented by the original position is an artificial device and knowingly thought to be so by the agents who employ it? If so, then it will be associated with the misrepresentation of a stable society in the way I have suggested in chapter one. That is, while the interpretation the agents have of their behind-the-veil reasoning sees it as merely artificial, they will represent their own thinking to themselves in a way that misconstrues what is required for a realistic representation of a stable society. While this response suggests that it might be possible to avoid the implications, for my thesis, of what Rawls says in his reply to Habermas it is unlikely that it can be so summarily dismissed.

For example, Rawls says elsewhere that, “citizens will of course differ as to which conceptions of political justice they think the most reasonable,” (Rawls, 1997, 770) and that “criticism of political liberalism as not allowing new and changing political conceptions of justice is incorrect” (ibid., 775, n.30). These quotes reinforce the view that Rawls envisages citizens in a well-ordered society as being capable of a perspective that embraces political and moral points of view, and their associated neutral and doctrinal perspectives. However, if this is what Rawls intends, his position is weakened by his

§5. The Artificial Defence

It might be thought that whatever the problem is said to be with Rawls' conception of the problem of stability it cannot be a genuine worry for Rawls. At the heart of the complaint is the view that Rawls defends a mistaken conception of reflective criticism. What needs to be tempered in a person who is obedient to institutions he or she would rather reject is the motive power of heartfelt concern. The way Rawls conceives reflective evaluation, however, makes obliteration of the

insistence, in the same paper, on the fundamental division between comprehensive and political outlooks. On this he says, for example,

Since the idea of public reason specifies at the deepest level the basic political values and specifies how the political relation is to be understood, those who believe that fundamental political questions should be decided by what they regard as the best reasons according to their own idea of the whole truth – including their religious or secular comprehensive doctrine – and not by reasons that might be shared by all citizens as free and equal, will of course reject the idea of public reason. (Rawls, 1997, 771).

It is imperative to realize that the idea of public reason does not apply to all political discussions of fundamental questions, but only to discussions of those questions in what I refer to as the public political forum. (ibid., 767).

These quotes suggest that in order to consider basic institutions from a neutral perspective agents must step outside of and in so doing detach themselves from the point of view associated with their doctrinal commitments. Indeed, Rawls refers to “what *Political Liberalism* calls the background culture of civil society in which the ideal of public reason does not apply.” (ibid., 775, n.28). The accusations I make in the text are, then, arguably supportable, I think, even in the face of evidence that Rawls wants to defend a conception of the agent that permits an ongoing critical stance towards the principles of justice and their rationale.

heartfelt the only option for one who, having engaged in reflective critique, accepts the judgement that goes against him. The motive power of the heartfelt can only be tempered from within a way of seeing that is informed by the commitments, convictions, affections, and so on that generate heartfelt concerns. A piece of rational reflection that does not engage those commitments, those affections, and so on, can result in a person overriding the motive power of heartfelt concern only by radically removing it. This fails to accommodate the tension between rational commitment to less than desirable institutions and the pull of personal preference, and this tension, I have argued, is required in an adequate representation of the problem of social stability. But Rawls has argued that the conception of the person in political liberalism ought to be viewed as artificial, as a theoretical construct, or heuristic device, intended to model but not to mirror actual evaluative capacities. Can't this same defence be made against the argument that he misconstrues the problem of stability? I want, in this section, to argue that this is not the case.

The defence that the conception of the person he employs is artificial arguably protects Rawls against the charge that he misrepresents personal evaluative capacities. The criticism I am making, however, does not point to Rawls' conception of the person *per se*, but to its implications for the representation of stability, and that representation, I will suggest, ought, by Rawls' own lights, to reflect something of the real world.

§5.1 Sandel and the Artificial Conception of the Person in Rawls

I begin by considering the way in which Rawls' appeal to the artificiality of his conception of the person in his theory seems an effective counter to the neo-Hegelian critique. I focus on the way it works as a response to Michael Sandel's objections in this regard. According to Sandel, Rawls' conception of the person has us,

stand to our circumstance always at a certain distance, conditioned to be sure, but part of us always antecedent to any conditions...Deontological liberalism supposes that we can, indeed must, understand ourselves as independent in this sense. (Sandel, 1998/1982, 11).

That is, we must understand the self as having an ontological priority that is the basis for our authority over the convictions we have, what permits us to view things from the value-neutral stance required by the theory. Sandel (1998/1982, 16) goes on to say why he thinks Rawls sees reflective, evaluative neutrality as a desirable characteristic. What he desires, says Sandel, is an "Archimedean point", a perspective from which societal arrangements can be assessed that does not locate the assessment within the existing modes of justification the society would use to defend its political arrangements. What is needed is a reflective space between the object of evaluation (the society's legal, politico-economic and social structure) and the means by which evaluation of the object proceeds (Sandel, 1998/1982, 16-17). What Rawls fails to do, on Sandel's account, is to provide a distance between the means of evaluation, and its object, that avoids an alternative pitfall. In placing the Archimedean point outside of existing modes of argumentation, Rawls asks the individuals in the society under inspection to adopt

a detached perspective. The methods of reasoning associated with this point of view must be arbitrarily selected, having no grounding in the evaluative commitments and ways of reasoning they regularly apply and see as, in fact, adequately grounded. Sandel says:

If the principles of justice are derived from the values or conceptions of the good current in the society, there is no assurance that the critical standpoint they provide is any more valid than the conceptions they would regulate, since, as a product of those values, justice would be subject to the same contingencies. The alternative would seem a standard somehow external to the values and interests prevailing in society. But...where the first would be arbitrary because contingent, the second would be arbitrary because groundless. Where justice derives from existing values, the standards of appraisal blur with the objects of appraisal and there is no sure way of picking out one from the other. Where justice is given by a priori principles, there is no sure way of connecting them up. (Sandel, 1998/1982, 17).

The way that Rawls squares this circle is, Sandel tells us, through his belief that the self is a distinct item that can reflectively regulate its ends. The individual has the capacity, on this view, to rise above context and to bring into critical view her convictions, beliefs, modes of evaluation, and the like, and subject them to rational scrutiny. While this conception does not require the noumenal, extra-natural, self of Kant's moral philosophy, and the difficulties that entails, it does, nonetheless, defend a similar role for reason in practical deliberation. The individual has a capacity for self-command that enables her to bypass extant desires, affections, and beliefs. Hence, Rawls, on Sandel's reading, is alive to and addresses the difficulty posed by the potential gap between a person's moral context and a reflective stance, and is convinced that his conception of the person,

as fundamentally under the command of a detachable self, is both correct, and serves to link what is worked out in reflection with everyday convictions. However, says Sandel, Rawls' conception is faulty in that it has the individual appeal to a normative basis for evaluative judgements that lie outside of her ordinary way of thinking. This alien basis can only be grafted onto and not integrated with an individual's ways of thinking and so be, for him, groundless. Sandel recognises that in *Political Liberalism* Rawls abandons the notion that his theory of justice depends upon a voluntaristic conception of the person - where "voluntaristic" captures the idea that a person is in command of her desires, beliefs, and so on.

It is in response to this that Rawls adjusts the interpretation he offers of his theory by having it float free of any particular, substantive, understanding of the relationship between the self and choice. His arguments are, he says, in this revised version, political; innocent of any view about the metaphysical status of the self. Hence, the nature of the self as such is not at issue. And so, as Sandel himself acknowledges, in Rawls' later work, "(t)he priority of the right over the good is not the application to politics of Kantian moral philosophy, but a practical response to the familiar fact that people in modern democratic societies typically disagree about the good." (Sandel, 1998, 189). A grounding for right that is neutral among conceptions of the good is a necessity given the fact of radical disagreement among people within a "constitutional democracy". Rawls recognises his earlier work to have been a project in "comprehensive liberalism", a liberalism founded on arguments intended to convince all irrespective of what they see as being of value in life. This comprehensive doctrine, associated in

intention with, as Rawls sees it, Kant and Mill, is replaced in *Political Liberalism* by a "political conception of justice" that does not expect to root in common soil an agreement about the basis for the just society. The source of agreement and the basis for cooperation is sought instead in an "overlapping consensus" (Sandel, 1998, 190). This does not, however, Sandel thinks, remove the need for a conception of the person. For the results churned out of the original position device to have any purchase on actual people the theory must presuppose something about them, to secure the possibility of their being responsive.

(In order for this way of thinking about justice to carry weight, the design of the original position must reflect something about the sort of persons we actually are, or would be in a just society." (Sandel, 1998, 191).

But Rawls insists that the parties to the original position are,

merely the artificial creatures inhabiting our device of representation. Justice as fairness is badly misunderstood if the deliberations of the parties, and the motives we attribute to them, are mistaken for an account of the moral psychology, either of actual persons or of citizens in a well-ordered society. (Rawls, 1996/1993, 28).

And this is a powerful rebuttal. Sandel has argued convincingly that Rawls cannot, without inviting a charge of arbitrariness, construct a reflective stance that fully severs the associated mode of evaluation from the person's ethical outlook. Nor can he embed rational critique in standing modes of thought. However, Rawls, equally convincingly, side-steps the issue by adopting, or emphasising, the non-metaphysical character of his theoretical *avatars*. The conception of the

person need not represent either real people or the, within-theory, citizens of a just society. In *Justice as Fairness: Political not Metaphysical* (1985), Rawls says,

As a device of representation the original position is likely to seem somewhat abstract and hence open to misunderstanding. The description of the parties may seem to presuppose some metaphysical conception of the person, for example, that the essential nature of persons is independent of and prior to their contingent attributes, including their final ends and attachments, and indeed, their character as a whole. But this is an illusion caused by not seeing the original position as a device of representation...when we simulate being in this position, our reasoning no more commits us to a metaphysical doctrine about the nature of the self than our playing a game like Monopoly commits us to thinking that we are landlords engaged in a desperate rivalry, winner take all. (1999, 402-3).

For Sandel, this does not get Rawls off the hook. Sandel's comeback, however, (in the second edition of his *Liberalism and the Limits of Justice*) does not rely centrally on his original metaphysical charge (see also Sandel (1996, 17-20)). The difficulty now, as Sandel sees it, is that there is nothing to prevent someone, some group or representative, adopting a reflective stance that keeps their value-related characteristics intact. That is, the non-metaphysical construction does not carry with it a reason for an individual to adopt the neutral stance Rawls requires to make his argument work. Sandel asks (1998, 191) "Why must we 'bracket', or set aside, our moral and religious convictions, our conceptions of the good life?" While Rawls distinguishes between moral choice in general and political choice in particular, and insists that "for political purposes...we should think of ourselves as free and independent citizens, unclaimed by prior duties or obligations." (Sandel,

1998, 192) this does not of itself supply a reason to be "unclaimed by prior duties" in our reasoning about political institutions.²⁷

Rawls' political conception, however, puts Sandel on the back foot. No longer able to press a metaphysical charge he objects to the disparity between the artificial construction and the agents it, he thinks, is supposed to represent. But when Rawls point out that the model agents are not intended to be representational Sandel's point seems to have little purchase. We can, with little difficulty, imagine an artificial representation of ourselves for purposes of argument without supposing that the model captures something true about our actual psychology. The economist's representation of the individual as a rational utility maximiser, for example, does not require us to think that it tracks our actual practical thinking for it to be of use. (Friedman, 1966, is the classic statement of this in economics). Sandel has other objections. He remarks, for example, that "the political life that (political liberalism) describes leaves little room for the kind of public deliberation necessary to test the plausibility of contending comprehensive moralities" (Sandel, 1998, 211); and that not taking into account one's own and others' moral and religious convictions in public discussion is only one way to establish mutual respect. Another way to show respect is to engage as far as possible with the perspectives of others, attempting to arrive at an understanding of their point of view by exploring patches of shared meaning, to

²⁷ Sandel (1996) traces an oft noted modern malaise in modern American life, a lack of connection between people and public institutions, to precisely the liberal neutrality Rawls recommends for the political point of view. Sandel's suggested fix is a revival of a republican conception of democracy, and the generation of civic virtue, grounded in an understanding of the common good.

grasp not only the strength of their beliefs but the qualitative aspects of their view of the good (This is often referred to as "Evenhandedness", see Carens (2000)). These may well be sound arguments, but they do not challenge, as Sandel's original critique did, the status of the conception of the person Rawls employs in his argument. A response that is favourable to but not mentioned by Sandel is as follows. In connection with the division of points of view referred to above Rawls says:

The third point of view – that of you and me – is that from which justice as fairness...is to be assessed. Here the test is of reflective equilibrium: how well the view as a whole articulates our more firm considered convictions of political justice, at all levels of generality, after due examination, once all adjustments have been made. (1993, 28).

Sandel could claim, I think, that the process of reflective equilibrium cannot plausibly come to rest unless a more realistic, situated, conception of the person is adopted.²⁸ Excluding a facet of our moral characters from the artificial conception – our situatedness – removes the possibility of a harmony between actual fact and theoretical device. A key ingredient, Sandel might argue, would always be missing. But I doubt that this argument can be made to work. A reflective equilibrium, a harmony of intuition and theoretical account can contain,

²⁸ (1996, 385) "Reflective equilibrium...is a point at infinity we can never reach, though we may get closer to it.." Given this specification of the method of reflective equilibrium, perhaps the way to put the point in the text is to say that, without an adequate denotational representation of the person, reflective equilibrium will be *asymptotic* due to a misspecification (of the person) rather than because of its character as a method.

unproblematically, a representation of the person designed to theorise but not represent actual psychologies.

§5.2 Stability in a "Realistic Utopia"

Rawls' rejection of Sandel's remarks about his reliance on the unencumbered self is, arguably, successful. While Kymlicka (1991) and Barry (1995), it seems, think that Rawls ought not to have seen Sandel's criticisms as so pressing as to require a re-examination of his theory, the emphasis on the artificiality of the conception of the person, at the centre of his work, serves Rawls well as a defensive strategy against Sandel and his like. Does it likewise serve to defend Rawls against an attack on his conception of stability? I would argue not. Consider the following quote from Samuel Freeman:

To appreciate the development of Rawls's views it is essential to understand that all along he has sought to work out a realistic ideal of justice (a "realistic utopia"). His conception is ideal insofar as it is designed for the ideal conditions of a "well-ordered society", where reasonable persons who are free and equal all accept the same conception of justice. Rawls's account of justice is realistic since it is designed to apply neither to moral saints or perfect altruists on the one hand, nor to natural sinners or rational egoists on the other, but to what humans at their best are capable of, given their nature, under normal conditions of social life." (Freeman 2003, 2).

Rawls formulates his project as one that aims to establish principles for a "realistic utopia" in his *Law of Peoples* (Rawls, 2001, 4, 5-6, 11-12). There he says that he is echoing Rousseau in his intention to take people as they are and "laws as they might be" (Rousseau (1950/1762, 3). But Rawls makes comments on the intended realism of his project elsewhere as well. He says in *Political Liberalism* (p.39),

when commenting on *Theory's* mistaken reliance on a common doctrinal commitment to public institutions, that his corrections, "do not impose the unrealistic – indeed, the utopian – requirement that all citizens affirm the same comprehensive doctrine.." (Cf. Rawls, 1996, 86-87).

So, the principles that regulate the well-ordered society could regulate real world politics, in the sense that the requirements they impose are within the range of those that actual people could live under, taking into account facts about how we respond to incentives, our, typically, limited altruism, and so on. Hence, we can distinguish between the conception of the person, or her reflective capabilities, and the contribution these make to the character of the well-ordered society. While Rawls might argue that the conception of the reflective person thought to be at the root of the weakness in the representation of stability is merely artificial, he cannot say the same about the representation of stability that the artificial conception supports. Hence, even if we accept the artificial character of personal reflection we can accuse Rawls of failing to respect his own standards of realism by failing to accommodate heartfelt concern in his conception of social stability. There is a criticality in real world democratic societies, a tension between compliance and critique that is absent from Rawls' well-ordered society. The well-ordered society is intended to be realistic in that it is attainable, but in the absence of criticality the well-ordered society is a long way from being a world that could be lived in.

§6. Conclusion

The neo-Hegelian critique of Rawls' theory of justice doubts its ability to accommodate the influence of culture on a person's evaluative reasoning. Two responses, one that describes the social world in such a way as to credibly combine cultural embeddedness and autonomous critique, and a second that emphasises the artificiality of the concept of the person in Rawls' theory, can be seen to fail when confronted with a criticism of Rawls' representation of the problem of stability. Neither overcomes the privileged position Rawls gives to the impersonal point of view and the difficulty this creates for the incorporation of the critical tension that is the mark of real world, stable but potentially unstable, democratic societies. This is significant for Rawls, I think, because stability is not a side-issue for him, but is a central motif throughout his work. In the chapters that follow I try to work out a response that draws on his abiding presumption that the well-ordered society is a market economy. In the next chapter I introduce some concepts, that will be drawn on in chapter three, through a consideration of an approach to Rawlsian theory associated with G. A. Cohen.

Chapter 2

The Embeddedness of Market-Oriented Choice:

A Response to Cohen on Rawls

Introduction

The purpose of this chapter is to prepare the way for the discussion of chapter three by formulating some important distinctions. Primarily, I want to emphasise the distinction, as it appears in Rawls, between the role of comprehensive doctrine in everyday, market choice, and the role of institutional restrictions, which perform a more regulatory function. That is, while, comprehensive doctrine informs personal choice, on each occasion of choice, institutional regulations place limits on what can be legitimate objects of choice. Institutional rules place prior limits on choice with regard to what a person is permitted to choose, but, unlike personal doctrine, do not inform deliberation regarding what, on each occasion of choice, what to choose. In order to bring out these distinctions I draw on a prominent strand in recent criticism of Rawls' theory of justice, one that contends that his theory is vitiated by a limitation he places on the reach of the principles of justice he defends. They ought, according to Rawls, to apply primarily to a society's basic institutions, it being a further question what principles ought to be taken to apply to personal conduct. According to those who oppose this gulf between institutional and personal principles, if principles are

proposed for institutions then they ought also to regulate personal choice in everyday social and economic contexts²⁹. The two main advocates of this "monist"³⁰ critique are G. A. Cohen (1992, 1995, 1997, 2000) and Liam Murphy (1998), for whom, "all fundamental normative principles that apply to the design of institutions apply also to the conduct of people." (Murphy, 1998, 252). So, in this chapter I discuss this area of Rawlsian criticism in order to establish some background understanding, essentially of the relationship between institutional and everyday choice in a market context, in Rawls, as preparation for the argument of chapter three.

Defenders of the view that Rawls is right to give principles of justice an institutional focus, are offered by Andrew Williams (1998) and, more recently, Thomas Pogge (2000, 2002a). For Cohen, on Pogge's reading, the objective,

²⁹ Cohen's claims apply to social as well as the type of economic conduct that is mediated through markets. Giving a lower priority to the educational needs of female members of the household damages their expectations and reduces their actual and potential holdings of "primary goods" in a way that violates the difference principle. My discussion applies only to conduct that can be straightforwardly categorised as market-based activity.

³⁰ Pogge (2000) suggests the "monist" tag is misplaced. We might argue, for instance, that principles for institutions entail principles for economic conduct while leaving the social sphere, for example internal family relations, untouched. That is, there may be more than two "sites" of distributive justice, and the term "monist" doesn't pick this up. Also, though not mentioned by Pogge, Rawlsian dualism commonly refers to the distinction between political commitments and the "thick" conceptions of the good life associated with personal doctrines. Adopting "monism" for a Cohen-Murphy position will not be a suitable opposition to dualism in this latter sense; as it need not pay explicit attention to the neutral/thick distinction. Nonetheless, I find it convenient to use the term "monist" to refer to the type of position defended by Murphy and Cohen.

embodied in principles of justice, that inspires institutional design should also inspire everyday personal decisions. On Pogge's reading, then, Cohen is claiming that Rawlsian citizens ought to modify their actions so as to directly promote the aims favoured by the difference principle.³¹ And Pogge offers criticisms that Cohen, read this way, would certainly find awkward to deal with. For example, if individual decisions are to be informed by the difference principle, in the form of a moral requirement, then,

(j)ust as people can erase socioeconomic inequality while raising the lowest socioeconomic position by giving some of their money to the poor, they can promote the same goal by expropriating others' money. And why should they not feel morally compelled to do so? (Pogge, 2000, 162).

However, Pogge works with an interpretation of Cohen that emphasises the need to look to personal actions and the possible outcomes of those actions when assessing their permissibility, a reading that is, I want to argue, not compulsory.

An alternative take on Cohen is available, one that provides a way to avoid Pogge's criticisms, while retaining his core claim against Rawls: that he ought not to see determining the details of a personal morality as a secondary task for a theory of justice. On the interpretation of Cohen I propose, Rawlsian citizens must look not to actions per se, nor to their consequences, but to the *reasons* for the choices they make. It is certain reasons for choice that are impermissible, not, primarily, certain acts.

³¹ This is not to say that they must direct their attention in particular choice scenarios to the least well off in that scenario (Pogge, 2000, 161, n.43).

In section one I give a more detailed account of Cohen's view, and go on to develop the reasons-based account by looking first at Williams' paper (section two), and then Pogge's critique (section 3). A retreat to a reasons-based account of Cohen's argument does not, however, make Cohen safe from criticism. I argue first (section four), that Cohen, along with Murphy, pays insufficient attention to the importance of the distinction between the neutral and non-neutral standpoints associated with, respectively, political and everyday concerns in Rawls' later work. Once this distinction is properly respected we see, I argue, that Cohen must (and will find it hard to) characterise personal principles simultaneously as principles that inform everyday personal deliberation *and* principles supported by the neutral reasoning appropriate to institution-level concerns. To properly respect the distinction in Rawls between neutral and non-neutral standpoints is to recognise that principles that inform institutional choice cannot, except perhaps unusually, have for their basis the convictions that constitute a comprehensive doctrine. If principles that influence institutional design and principles that inform everyday personal choice do, in the usual case, share a common basis in personal ethical outlooks, then the division Rawls sees as so important in his later theory between political commitments and commitments shaped by comprehensive doctrine is threatened. Applying this to market activity (section five) the question arises how Cohen can characterise principles for personal market-based choice as at once part of everyday practical reasoning and entailed by reasoning associated with a neutral stance. We might in response to this want to wipe out the role of comprehensive doctrines in personal market choice, but there is, as I try to show in section five, good reason to think that this cannot be done without damaging

(an important) part of Rawls' thesis. Finally (section six), I consider a response that is available to Cohen. He can claim that the personal principles he sees as entailed by the difference principle are, what I call, *delimitative*. That is, he can claim that we need not see such principles as operative *in* personal deliberation, but as delimiting, a priori, what a person can permissibly take into account in personal choice. In this way they will not interfere with the bases of comprehensive doctrines but will, Cohen might claim, nonetheless inform personal choice. This, I suggest, will not work. A delimitative principle that places parameters on choice but does not play a role *in* practical deliberation is not a personal principle, a principle that informs everyday choice. It doesn't feature in first-personal, *in situ* practical reasoning, and so fails to place requirements on choice in the way Cohen needs in order to make his thesis compelling. Section seven concludes.

§1. Cohen's thesis

In the central economic aspect of his argument, Cohen criticises Rawls for not applying his difference principle "in censure of the self-seeking choices of high-flying marketeers, which induce an inequality that...is harmful to the badly off" (Cohen, 1997, 5). They ought, Cohen thinks, to eschew material rewards as an incentive for taking up more demanding economic and social roles. In general, according to Cohen,

principles of distributive justice, principles, that is, about the just distribution of benefits and burdens in society, apply wherever else they do, to people's legally unconstrained choices. Those principles...apply to choices that people make within the legally coercive

structures to which, so everyone would agree, principles of justice (also) apply.' (Cohen, 1997, 3).

And, as applied to Rawls, in particular, the claim is (in its economic aspect) that the difference principle, Rawls' principle of distributive justice, ought to inform everyday market-based deliberations as well as institutional arrangements. Cohen's central argument for his position as it applies to Rawls can be put as follows.

Rawls defends two theses:

Thesis A: A society whose members accept the difference principle can take it to regulate background institutions while not taking it to inform their personal choices. Rawls says,

The difference principle holds, for example, for income and property taxation, for fiscal and economic policy. It applies to the announced system of public law and statutes and not to particular transactions or distributions, nor to the decisions of individuals and associations, but rather to the institutional background against which these transactions and decisions take place. (Rawls, 1993, 283).

Thesis B: A commitment to the difference principle on the part of the agents within a Rawlsian society is compatible with their unrestrained pursuit of material rewards intended to foster higher productivity. As Cohen puts it:

It is commonly thought, for example by Rawls, that the difference principle licences an argument for inequality which centres on the device of material incentives. The idea is that talented people will produce more than they otherwise would if, and only if, they are

paid more than an ordinary wage, and some of the extra that they will then produce can be recruited on behalf of the worst off. (Cohen, 1997, 6).

Thesis B is, argues Cohen, indefensible, for the following reason. It cannot be maintained that extra reward is necessary to entice additional work effort from the talented in any but a casual (non-logical, non-causal) sense of the term "necessary". Indeed, extra reward is necessary only because the talented will not agree to work harder (or more efficiently) without it. Their not agreeing is not for them a matter of, say, invincible psychological fact, but a reluctance on their part that they could with effort overcome (For example, Cohen, 1997, 6-10; 2000, 121-122).³² Hence, expansion of social output through enhanced productivity could be forthcoming without additional productivity incentive. This inequality, then, is unjustified by the difference principle the Rawlsian individuals avowedly accept.

This implies that the difference principle entails a restriction on personal choice - and this entailment dissolves the divide Rawls thinks he can defend in thesis A.

³² It is interesting to note here one of the senses Aristotle attaches to the term "necessary". The quote is from Wiggins (1998/1987, 25): "Aristotle is not usually credited with having any view of what a need is; but in his philosophical lexicon at *Metaphysics V* he has the following entry for 'necessary': 1015a20: We call NECESSARY...(b) anything without which it is not possible for good to exist or come to be, or for bad to be discarded or got rid of, as for instance drinking medicine is necessary so as not to be ill, and sailing to Aegina so as to get money.." Cohen would argue, I think, that incentive payments for greater productivity, on the part of the talented, are not necessary in Aristotle's sense b.

§2. Questioning Williams' ethos-based reading of Cohen

Williams (1998) presents an account of Cohen that, in its essentials, chimes with that presented above. In the context of market choice, advantaged, talented people ought to forego material rewards intended to provide incentives for extra effort. Such effort could be made in the absence of such incentives and so the inequality they occasion cannot be classed as *necessary* in the sense required for them to be justified by the difference principle (see Williams, 1998, 226-227). However, Williams' response to Cohen relies upon a reading of him that emphasises the role of ethos in his argument. That is (according to Cohen on Williams' reading) Rawls ought also to recognise the need for a social ethos, that embodies moral requirements on personal conduct, in order to promote compliance with the personal principles Cohen argues for. There is, as one might expect, ample evidence that Cohen does think that "the difference principle requires an egalitarian ethos." (Williams, 1998, 227)), for there to be widespread compliance. On the one hand, Cohen does not envisage a community of moral saints:

...in estimating what it would be like for a person to accept a salary that is much lower than what full exercise of market power would provide, the strain to think about is the one he would feel if, *ex hypothesi*, people like him are accepting similarly modest salaries. We are talking about an egalitarian society, not about a population of talented people each of whom is a unique moral hero. (Cohen, 1991, 293).

On the other hand, government cannot be relied upon to ensure compliance with personal moral principles in the marketplace (Cohen, 1991, 315). Hence, given a recognition on the part of Rawlsian agents that they cannot defend their "need" for productivity payments, an egalitarian ethos is required to foster (what need not

be self-conscious (p. 316)) personal compliance. Williams' response to Cohen makes the latter's remarks on the necessity of an egalitarian ethos a prominent feature in his argument, and, on the basis of this interpretation, Williams' argument is a powerful one.

The rules that constitute institutions satisfy, Williams points out, a "publicity" requirement Rawls imposes on principles of justice, a requirement, centrally, that violations of them be transparent (Williams, 1998, 233). The basis of a decision made in the marketplace, including a decision concerning career choice, or remunerative package is not so transparent, and this lack of transparency - or inherent asymmetry of information between a market actor and those who would judge him - makes market choice a private and not a public affair for the purposes of Rawlsian justice.

To be more precise, Williams offers a strengthened version of the *basic structure objection* to Cohen's argument. According to this objection the difference principle is "inherently restricted" (Williams, 1998, 228) having application only to the basic structure and not to personal conduct. Williams notes examples of the textual support for such a restriction in Rawls and considers Cohen's main response. Cohen argues that market conduct ought not be seen as "occurring within, rather than comprising part of, a society's basic structure." (Williams, 1998, 229). If, on the one hand, Cohen argues, basic institutions are restricted to being those that are "legally coercive" (Williams, 1998, 230) then it is hard to see how this restriction can be defended. Principles that govern institutions in this sense will be silent about morally repugnant conduct merely because it lies outside the range of conduct that is subject to legal coercion. If, on

the other hand, basic institutions are understood to be those that “produce profound effects” (Williams, 1998, 231) then, argues Cohen, the behaviour of market participants in their pursuit of material rewards ought to be seen as part (though an informal part (ibid.)) of the basic structure, and this because of its profound effect on the opportunities facing others.

Given the strength of Cohen’s response to the *basic structure objection* Williams restates it in order to strengthen its appeal. To do this he notes that “the basic structure can be defined either by reference to what it does (its *dispositional* properties), or how it does whatever it does (its *intrinsic properties*)” (Williams, 1998, 231. Phrases in parentheses not in original). Williams then observes an intrinsic property overlooked by Cohen that secures a restriction on the reach of Rawls’ principles of justice that excludes the decision to be a market-maximiser. The intrinsic property of basic institutions Williams describes is the public nature of their constitutive rules (Williams, 1995, 233):

Rawls appears to regard institutions’ constitutive rules as public in three respects. Thus individuals are able to attain common knowledge of the rules’ (i) general applicability, (ii) their particular requirements, and (iii) the extent to which individuals conform with those requirements.

The decision whether to be a market-maximiser is not public in the sense Williams associates with the basic structure. We can see here that the non-public character of market conduct is most evident in connection with point iii. The lack of transparency of motives and rationales for wage offers would undermine any attempt to publicise the extent to which individuals were conforming with a principle such as that which Cohen recommends.

Williams recognises that it is important in order for his argument to work that the ethos Cohen argues for “cannot be regarded as realizing public rules.” (Williams, 1998, 235) and goes on to show that it would be difficult for Cohen to maintain such a position.

I want to argue, however, that there is an alternative reading of Cohen that retains the force of his criticism against Rawls while making the role of an egalitarian ethos, and its status, a secondary rather than a central issue. If this can be sustained then Williams' defence of Rawls may not be quite so compelling. The insistence that Williams points to in Cohen on the necessity of an egalitarian ethos suggests that we ought to picture the Rawlsian agent as responsive to the personal principle he (Cohen) recommends, only to the extent that it is a principle embodied in the requisite ethos. In order to argue for an alternative construal of Cohen's opinion, I want, first, to try to show, against this, that the personal obligation Cohen points up can be seen as pressing for a Rawlsian agent irrespective of the existence of an egalitarian ethos. That is, the (talented) Rawlsian can, indeed ought, to see that he has reason to accept what Cohen says by assessing the compatibility of his commitment to the difference principle with his willingness to accept, or determination to push for, productivity incentives. There is much in Cohen to prevent this sort of reading. First, Cohen appears to argue that pure productivity payments³³ are not, as such, morally indefensible, even where they are accepted by a supporter of the difference principle. Though a person would, in Cohen's opinion, be taking an implausible, mechanistic view of

³³ I call them "pure" to distinguish them from (by Cohen's lights legitimate) payments made, for example, to compensate for arduous or risky jobs.

the behaviour of the talented, a person could argue that material incentives just are necessary to get the talented to exert themselves. He would not then be making a moral mistake. His mistake would be a factual one. His (mistaken) belief in the will-independence of talented persons' need for pure incentive payments would be compatible with his commitment to an egalitarian principle that sanctions only inequalities that are necessary to enhance the position of the less well off. That a person who defends differentials due to pure productivity payments is morally at fault becomes apparent only when we imagine her (a talented person) articulating a first-personal defence of pure productivity payments (Cohen, 1991, esp. 271-304). It is apparent in this voicing of her view that the necessity she ascribes to the talented she is ascribing to herself, and there is no room here for factual error - no distance between speaker and subject matter that might justify a misunderstanding of the relationship between possible effort levels and the rewards necessary to entice them. In making the argument, first personally, the speaker reveals her "inegalitarian attitude" (Cohen, 1991, 269). The argument that incentives are required to expand the social product, and so enhance the position of the least well off, is successful only if she thinks it justified, for her and people similarly talented, to *refuse* to put in extra effort without additional reward. But such refusal is incompatible with a commitment to the difference principle, a principle that sanctions only inequalities that are necessary for the betterment of the less well off. It is not merely a first-personal account, however, that reveals the injustice here. It also has to be the case, Cohen argues, that the speaker recognises herself to be part of a "justificatory community", for her acceptance of productivity

differentials to be morally criticised; and a justificatory community, Cohen tells us (1991, 282) is,

..a set of people among whom there prevails a norm.....of comprehensive justification. If what certain people are disposed to do when a policy is in force is part of the justification of that policy, it is considered appropriate to ask them to justify the relevant behaviour, and it detracts from justificatory community when they cannot do so.

Indeed (Cohen, 1991, 307) the,

incentive argument is not problematic.....when it is thought acceptable to view the rich as outside the (justificatory) community to which the poor belong.

Making his point depend upon the assumption that Rawlsian society forms a justificatory community is surely innocuous, but it does imply that his argument, for him, only has force if we picture the talented actually justifying their action with regards to productivity payments to the less well off. This seems to cut off the option of presenting a monological alternative to Williams' ethos-based reading. The talented must address their commitments and behaviour post-contract not merely because they are enlightened by Cohen's argument, but because, it seems, they are also, and they recognise themselves to be, members of a justificatory community: "a set of people among whom there prevails a norm...of comprehensive justification." (Cohen, 1991, 282). But this need not prevent the reading I want to offer, one that avoids the emphasis on egalitarian ethos found in Williams' work.

It is possible for the talented person to reason that *were she to be* confronted with a less well off person, then, given that they are (by her own account) members of a justificatory community, in Cohen's sense, she would reveal her "inegalitarian attitude" in requiring, what are in fact, unnecessary incentives for additional personal effort. The revelation here is to the talented individual herself, rather than a public shaming. This monological alternative lays emphasis on a talented person's beliefs about what is permissible personal conduct, given a commitment to the difference principle. This account opens up the possibility of retaining the important claim against Rawls, that a commitment the difference principle entails a moral constraint on personal conduct, while making the question of how that recognition might be registered in actual personal decision a further question. It seems correct to think, with Cohen, that an ethos would be necessary, but this is, once we emphasise the monological character of the argument against the Rawlsian marketeer, a secondary consideration; one that takes into account the prevalence of weakness of will and the limits to effective government regulation.

Williams' argument depends for its cogency on the fundamental importance to Cohen of an egalitarian ethos. However, Cohen need not, I have so far argued, make his argument the way he does. He can make the primary target the Rawlsian agent himself, requiring him to amend his beliefs about what is permissible personal conduct, given a commitment to the difference principle. That an ethos, rather than iron will or governmental intervention is the only feasible way to have agents successfully exercise self-restraint is then a secondary point. The primary

point is that each has a personal moral principle thrust upon him by a commitment to the difference principle.

§3. Pogge's outcome-oriented reading and a reasons-based alternative

Pogge's argument draws on a distinction between what he calls mastergoal and supergoal monism. The latter is, Pogge says (2000, 156), Cohen's favoured conception while the jury is out as regards which one Murphy would say he prefers. Pogge defends Rawls against both. Offering definitions, Pogge says that mastergoal monism,

postulates that there is one goal (or system of goals) of ultimate moral importance and that all social institutions and personal dispositions, ethos, and conventions, ought to be shaped so that they together optimally promote this common mastergoal. Persons, in particular, ought to have whatever conduct principles and dispositions in fact cause them best to promote this goal. (Pogge, 2000, 155).

The supergoal variant, alternatively,

postulates that there is one goal (or system of goals), which ought to be all-pervasive: exalted and prescribed in conventions and institutional rules, permeating the ethos, inspiring individuals. Persons, in particular, ought consciously to pursue (aim at, strive for) this goal. (Pogge, 2000, 155).

The difference, it appears, between these two variants of monism is that a mastergoal thesis is compatible with individuals not aiming in their practical decisions to achieve the mastergoal directly or consciously. Thus, a polity based on Richard Hare's indirect, two-level utilitarianism would be mastergoal monistic.

A single principle, that preference satisfaction ought to be maximised, permeates both institutional design and individual practice, however, individuals in Hare's theory can, indeed will typically, have aims that are not explicitly preference maximising. (Hare (1952, 1962, esp. 1981)). For an account to be supergoal monist individual decision must take for one of its guiding principles, the (or a) principle that informs institutional choice. For example, in Kant's theory (Kant (1998/1785)) no moral standing is given to the instrumental value of acting from motives other than duty. If one wants, then, action in the personal domain to have moral standing, actions at this level must be 'from duty'. They cannot, that is, as in the mastergoal thesis, be of moral worth simply because of their contribution to the master goal. Supergoal monism as its name and Pogge's definition suggest is goal-oriented. Where it differs from a mastergoal version (with respect to individual conduct) is in its requirement that individuals act, consciously, in a way that they think promotes the type of outcome that their institutional arrangements are intended to promote. In what follows I refer to this goal-oriented reading as "outcome-oriented". This works, as we will see, as a better description of Pogge's interpretation of Cohen, given what he says in the objections he raises. Additionally, I should say that the term "outcome-oriented" relates to what Pogge calls supergoal monism in that it is the consistency of the possible outcomes of personal action with society's goals that influences, according to supergoal monism, personal choice. Supergoal monism, then, requires that the individual behave in a way that reflects the social goals embodied in the principles that influence institutional design. The principles that shape, for example, economic policy ought also to be brought to bear in personal attitudes and everyday

decisions. Moreover, in order to reflect, in the personal domain, their society's goals, individuals must assess their behaviour, and the choices they make for their contribution to those overarching goals.

That this is the outcome-oriented understanding of Cohen that Pogge works with shows up in a number of places in his discussion. Pogge says of talented Cohenists (Pogge, 2000, 142)³⁴ that they, "should be ready to do managerial work at a lower reward multiple *so as to* increase the laboring net pay" (emphasis added). The phrase "so as to", here, betrays an emphasis on the outcome of a person's behaviour as a way to incorporate the difference principle in everyday life.³⁵ But, drawing on the remarks I made in connection with Williams' argument, a requirement to have an outcome-orientation in one's thinking, is a secondary consideration. A general requirement on the talented to take on work that calls for their special talents follows, in the argument I have offered, from a recognition of the incompatibility of a commitment to the difference principle and unnecessary demands for extra pay. Where it is only an enticement to greater productivity that is the source of the high pay for managerial work, then the enlightened Rawlsian, enlightened by Cohen's reasoning, will see reason to refuse the extra pay, and do the work anyway. But the reason is related to the

³⁴ In the example he is using here managerial work is the more productive job that only the "talented" are capable of. A "talented Cohenist" is, here, a Rawlsian agent who is convinced by Cohen's argument and wants to adjust his behaviour accordingly.

³⁵ For further evidence that Pogge's is a goal-oriented interpretation of Cohen consider Pogge's view that, according to Cohen, Rawlsian individuals "should want neither the basic structure nor their own personal conduct to produce inequalities that are not to the greatest benefit of the lowest socioeconomic position." (Pogge, 2000, 141).

incompatibility of the difference principle and the reasoning behind certain personal choices. There is no requirement to “do managerial work at a lower reward multiple *so as to*”, or in order to effect more equal rewards. Note also that Cohen says,

one function of the egalitarian ethos is to make conscious focus on the worst off unnecessary. What rather happens is that people internalize, and – in the normal case – they unreflectively live by, principles which restrain the pursuit of self-interest and whose point is that the less fortunate gain when conduct is directed by them. (Cohen, 1992, 384, as cited in Murphy – note omitted – (1998)).

A reading, such as Pogge’s, that sees his core thesis as involving the self-conscious application of the difference principle does not sit well with this comment on one of the purposes of an egalitarian ethos. It is, Cohen thinks, perfectly compatible with his thesis to have a self-effacing ethos in place that achieves the conduct he sees as required for consistency in Rawls’ theory. There might be no room, in such a society, for the “so as to” of intentional action Pogge envisions.

One way to support the self-effacing ethos Cohen appears to have in mind, is to envisage the parties to an agreement about how to design the institutions of their society recognising that they ought, for consistency, to have the difference principle inform their everyday lives. They might then go on to agree that a self-effacing ethos would be the best way to achieve universal compliance. (There are reasons, of course, to reject the idea that Rawlsian citizens could be self-effacing in this way. Williams (1998) has one connected with publicity at page 234). What they could not do is decide, in line with Pogge’s interpretation of Cohen, that the

difference principle will inform their personal lives as a reflective endorsement of their commitment to the worst off, and then suppose an ethos that precludes reflective endorsement. A view, then, that sees Cohen's primary point as being about a requirement on personal conduct, and only secondarily about how compliance is to be achieved, accommodates, in a way Pogge's emphasis does not, the role Cohen countenances for ethos.

Further, Pogge says (2000, 149), that,

If managing and laboring are objectively equally burdensome, then managers should be asked to redirect all special rewards attached to their work toward raising the lowest net pay rate. This is what, Cohen holds, a true commitment to the difference principle requires.

But suppose that a person, on the reading I am recommending, believes what Cohen wants him to believe, then he will think that requesting, or demanding, more pay in order to call forth extra productive energies is impermissible. He will also think that a society in which "special rewards" paid for extra effort are the norm is not a just society. However, he may find himself in a society in which extra pay is offered on the understanding that it is required in order to entice people into occupations that require the exercise of rare talents. What then is a person to do with any extra reward that accrues? He can refuse it or redistribute it, but nothing in the realignment of beliefs that Cohen's core argument calls for requires him to do these things with extra reward so gained. For one thing, this would involve a fruitless exercise on the part of an individual to change society through individual acts of giving. If, on the other hand, such an understanding does not permeate the market economy - an understanding that premiums for

effort on the part of the talented are needed – then a supposed requirement to refuse or donate extra reward, received for additional effort, is otiose. Such an economy might be the result of an ideal theoretic behind-the-veil decision to have all market actors (including producers) conduct themselves on the understanding that pecuniary motives of the relevant type will not operate.

The foregoing appears to suggest that accepting, and keeping, productivity premiums offered in the marketplace is consistent with a commitment to the impermissibility of *requiring* non-necessary rewards. And I think this is a plausible corollary to the reading of Cohen I am recommending. However, there is a relationship between, on the one hand, a belief about the incompatibility of the difference principle and personally requiring incentives for effort, and, on the other hand, actual action. Were a person to take a premium for effort *because* it offered personal gain for additional effort then the reason she is prepared to (and, indeed, does) act from would be incompatible with her professed beliefs. This does put a block on action – action done for a particular reason – but this restriction antecedes any consideration of the outcome of the action, in terms of the actual or possible resultant distribution. It is a restriction on action, that is, that makes no room for Pogge’s outcome-oriented reading. On this “reasons-based” account of Cohen’s thesis, then, a redirection of “special rewards”, just because they are rewards designed to promote greater productivity, is not required. It is enough for an individual to address his or her beliefs and to act only from reasons compatible with them in order to answer Cohen’s call. That extra pay is forthcoming might be a fact about a market economy, a fact that follows from

employers' assumption that individuals respond in a predictable way to pecuniary motives; a fact not easily, directly addressed by the individual in the marketplace.

If extra pay is given (because, for example, accepting Pogge's account, individuals with a lower "respio"³⁶ benefit from the higher pay paid to attract people with higher respios) we can take it to be the individual's prerogative what he does with it. Cohen's extension of principles for institutions to everyday life does not impose requirements on what individuals do with their wage packets.

Pogge's critique of Cohen relies on this outcome-oriented interpretation and what it implies for individual action. If the interpretation is non-compulsory, however, then Pogge's charges are similarly open to rejection.

The following points reinforce the view that Pogge's remarks are misplaced on an alternative, reasons-based reading of Cohen:

(1) Included on the list of Pogge's suggested specifications of Cohen's thesis is the following:

One might specify the moral duty Cohen seeks to inject into Rawls' theory as a duty to redirect the economic-rent portion of any special rewards one receives toward raising the lowest pay rate. (Pogge, 2000, 147).

³⁶ A person's respio is her "reservation net pay rate ratio...defined as the lowest reward multiple that would (barely) suffice to induce her to work as a manager." (Pogge, 2000, 145). The respio differs from the economist's more standard reservation wage. The latter is the lowest amount of pay a person is prepared to work for, while the respio takes into account the ratio of this amount to the pay available in another job. The difference between the two becomes important, for example, when there are a number of jobs against which a respio for a particular job is calculated (as on p.146). The difference is also in play in point 2 of section 3d below.

This, connected to the notion that more reluctant individuals receive less economic rent³⁷, entails that the less reluctant (i. e. less reluctant to take productive employment that engages their talents) have more of a distributive burden than the more reluctant. In fact the more reluctant an individual is the less she must transfer of her income, as less will be accounted for by economic rent. But this rewards the very reluctance Cohen finds objectionable.

But restricting Cohen to comments about personal consistency and the sort of reason a Rawlsian can with consistency act from makes Pogge's suggested specification inapplicable.

(2) Consider also an objection Pogge (pp. 149-150) takes to be related to one offered by Williams (1998, 235-37). Suppose (in order to represent Pogge's argument) a two individual society. One person is "talented" and one is not. A more productive job is available to the talented individual, paying £10 per hour, while the less productive job pays £5 per hour. The £5 difference is pure productivity incentive. The talented person can (and Pogge takes this to be her duty according to Cohen) take the more productive job and give £5 to be distributed evenly between them, leaving both with £7.50 each (per hour). Alternatively, the more talented individual could take labouring work at the rate of £5 per hour. She would, in fact, prefer to do this (Pogge has her say) because she is attracted to managerial work only at a wage multiple greater than 1. (If she

³⁷ Because the economic rent a person receives is the difference between her actual wage and her "reservation wage", the minimum she is prepared to work for, the more reluctant a person is (or the higher her reservation wage) the lower is her economic rent.

goes into more productive employment and transfers income to achieve equality the wage multiple is $7.5/7.5 = 1$ removing the incentive for her to take up the managerial post, given her respio (I note what Pogge means by “respio” at n.24, above). If she works as a labourer, however, incomes are only £5 each. Letting $(£7.50, £7.50) = D1$ and $(£5, £5) = D2$, $D2$ is less preferred to $D1$ by the difference principle. So long as an individual has a respio above 1 it will be rational for her to work as a labourer. Working in more productive employment and transferring income until the wage ratio is 1:1 takes the wage multiple below that which she requires as incentive.

Unless, then, Cohen requires individuals to take up more productive posts then, says Pogge, all with respio above 1 will retreat, as this individual did, to less productive work, producing outcomes that are ranked lower by the difference principle than those that would otherwise obtain. Pogge goes on:

Cohen in fact holds that a true commitment to the difference principle does not permit (a) retreat to labouring. This is expressed in all three of Cohen's essays, most clearly in (the Pareto Argument for Inequality), where he argues that the same enlarged social product Rawls wants to see produced thanks to the incentives that an unequal distribution ($D2$) provides should instead be produced by citizens personally committed to justice and distributed equally ($D3$). Clearly the social product would not be the same if (individuals) followed (their) preference for laboring work under $D3$. (Pogge, 2000, 150).³⁸

³⁸ Assumed here is that a person will not look to the absolute level of reward available from being a manager once the rewards are distributed (i.e., £7.50), but to the reward relative to that available as a labourer. Given that the respio indicates a person's reluctance to do one job rather than another it seems a reasonable assumption. It should be noted also that Pogge (again, reasonably) assumes that pay differentials that reflect the reluctance that the respio represents are, for Cohen, permissible.

And this suggests to Pogge a fourth reading of Cohen's intention:

Here is then the fourth interpretation of the moral duty Cohen seeks to add: The talented ought to contribute as they would in the society Rawls envisions even while receiving merely an equal share of the social product. (ibid.).

Cohen is, I suggest, on the reasons-based reading, unmoved by this complaint. The individual does enough to satisfy Cohen if she acts for reasons consonant with consistent commitments. Not acting in order to acquire additional material reward is sufficient to make her market-oriented practical reasoning consistent with a commitment to the difference principle. If the market institutions are such that higher reward (as pure incentive) is forthcoming, is attached to the sort of job her talents qualify her for, then she can, if she wishes, distribute it among her neighbours or make a special payment to the tax authorities. This latter largesse is not, however, required by Cohen on the restricted reading I am recommending.

Similarly, a Rawlsian who has addressed his commitments and is now disposed to have his convictions inform actual choices does not have to register a substantive requirement in his deliberation to choose one line of work over another. Cohen says, according to Pogge's paraphrase (quoted above), that the "same enlarged social product Rawls wants to see produced thanks to the incentives that an unequal distribution (D2) provides should instead be produced by citizens personally committed to justice and distributed equally." This does not have to mean that the same enlarged product must be forthcoming in this way, so entailing a requirement on the individuals to produce it. It can mean that any enlarged product that may or does occur ought not to be the outcome of incentive

seeking choice behaviour. This leaves it open whether any such enlargement will, in fact, be forthcoming, it being left to individuals to decide what employment they take up (once their commitments have been laundered of their inconsistency).³⁹

I suggest, then, that Pogge's critique of Cohen assumes an outcome-oriented reading to be correct. On this reading Cohen requires agents look to the effects of their choices in order to check their consistency with the objectives of the difference principle. Cohen need not, however, go so far as to require individuals attend to the compatibility of their actions with a commitment to the difference principle; they need only assess the consistency of their beliefs and, accordingly, restrict the reasons they can permissibly act from in market-based choice. Once this is recognised Pogge's criticisms miss their intended target. Redirection of a portion of the pay packet, and widespread embezzlement do not follow from a requirement to flush out inconsistent beliefs and certain motives of personal interest. Furthermore, a *requirement to take up* more productive work that engages scarce talents is not entailed in Cohen. Were Cohen to present, in fact, a thesis that has the shape Pogge takes it to have then his critique would, I think, be a considerable blow to Cohen's view. The same can be said of Williams' critique. What Cohen can do, I have argued, is to restrict his thesis to one that points up what Rawlsian agents' must for consistency believe to be legitimate reasons in personal, market-based decision.

³⁹ Similarly, the objections made in Pogge's section X depend entirely, I think, on an outcome-based interpretation of Cohen.

I present in the next sections a further challenge to the monistic thesis, as it applies to Rawls, to which the reasons-based version of Cohen's thesis is, however, not resistant. I begin by considering a challenge to the monist based around the importance later Rawls attaches to the distinction between neutral and engaged reasoning, the type of reasoning associated, respectively, with what Rawls marks out as political discussion concerning institutional choice, and reasoning that engages comprehensive beliefs and commitments, in everyday contexts.

§4. Rejecting the reasons-based version of Cohen's view

There are two aspects to my rejection of the reasons-based version of Cohen's thesis. The first aspect is an emphasis on the importance of neutrality in the derivation of institutional principles. The second aspect emphasises the social embeddedness of market conduct. Taken together they help to emphasise the divide between the institutional and the personal context of market choice. Market choice is at once influenced by institutional constraints, such as restrictions on the quantities of a good that can be traded, and by personal, for example religious and moral, concerns. This divide is important in my attempt in the next chapter to derive a conception of embedded critique by looking at the characteristics of the market participant.

§4.1 First aspect of the argument: The neutral basis of institutional principles versus the doctrinal basis of personal market choice

Cohen and Murphy are brief in their remarks on the difference between Rawls' earlier and later work, and its implications for their theses. In "Political Liberalism" Rawls attaches great importance to this difference. As noted above, Rawls came to see that requiring people to embrace the two principles of justice as a moral doctrine fails to respect the fundamental role of their personal moral doctrines in shaping their personal outlooks. The Rawls of *A Theory of Justice* took the device he used, to derive principles for just societal arrangements, to reflect fundamental moral convictions that all in a pluralist society could share. His later revision reflects his coming to see that visions of the good life, and of what constitutes fair social arrangements in the personal sphere, emanate from fundamental moral, philosophical, and religious convictions that differentiate social groups. The neutrality of the original position device denies the primacy of thick conceptions of the good, and the formative role they play in shaping the normative appearance of the world. This emphasis on the basic place of personal moral, religious, and philosophical beliefs seems to provide, alone, a sound basis for a dualistic conception of principles of justice. If principles for personal conduct derive from fundamental moral convictions, then it is inappropriate to have the same principles inform both institutional design and personal behaviour. If they inform what people do in their personal lives then they are coeval with their fundamental convictions. To suppose them also principles for institutional design bases political agreement on a supposed common ground, and this reinstates the errors of *A Theory of Justice*.

This creates problems for the reasons-based version of Cohen. On the reasons-based reading the agent is required to address the consistency of his

beliefs and manage his practical reasons accordingly. He is to eliminate as a possible reason for taking a job the premium offered for effort, effort that could be made without additional reward. What is required, on the part of the market participant, is, then, principled self-restraint. This requires that the principle be understood to be part of a person's background convictions. Were we to imagine it as a principle that emanates from outside of an individual's standing convictions we would be presenting a conception of principled self-restraint in which self-control is exercised, so to speak, "from behind the scenes", from outside of the person's moral outlook; and that invites an obvious incoherence, one familiar from the neo-Hegelian objections to Rawls' theory. We can, of course, construct such a picture. Kant, on a common reading, offers a conception of practical decision in which moral reasons bypass natural psychology. And, this can, in familiar ways, be countered with a rejection of the notion of there being such a thing as practical reason (For example, Blackburn (2001), Korsgaard (1986)). But this debate is not to the point. Whatever one's view about the capacity of reason to autonomously direct the will, what Rawls must avoid is not a position in this dispute. To satisfy the strictures of his later theory, a principle that entails self-restraint must either have a supporting rationale that appeals to or has its basis in standing fundamental convictions, or it must have a rationale that does not. If the former, it cannot be a principle also for institutional design as it has an improper provenance in existing commitments. But if its supporting rationale does not make appeal to moral convictions then it cannot form part of personal reasoning about what to do in everyday choice scenarios.

So Cohen, on the reasons-based reading, is faced with a dilemma. Either he supposes that the personal principle⁴⁰ individuals must apply has a rationale that lies outside of their background convictions or he supposes it to emanate from within existing moral doctrine. If the former it cannot have purchase in everyday life, if the latter it cannot be a principle also for institutional design. What makes this an acutely problematic dilemma for Cohen is that it cannot be overcome by assuming away the division between the neutral reflective stance and the embeddedness of personal doctrines. If he assumes this away then he does not have Rawls in his sights. Unlike (what according to Rawls are) comprehensive liberal positions such as Kant's or Mill's it is of central theoretical importance in Rawls that that divide be respected.⁴¹

⁴⁰ To be clear about the content of this principle: According to Cohen, for consistency, a Rawlsian agent must believe that it is impermissible for her to seek extra reward for additional effort that could anyway be offered. The personal principle that follows from this is, on the reasons-based reading, that a person ought not to let that part of any premium offered that is designed to entice extra effort influence the decision about whether to take a post or not. (As I have argued, the rejection or donation of extra productivity premia actually received is not entailed by this principle).

⁴¹ Recalling footnote 26, above, in *Political Liberalism: Reply to Habermas* (1995) Rawls presents an account of his theory that suggests the divide, as I draw it, between comprehensive and neutral thinking does not well represent Rawls' thinking. In his reply Rawls states that, "We assume that each citizen affirms both a political conception and a comprehensive doctrine..In this case, the citizen accepts a political conception and fills out its justification by embedding in it some way to the citizen's comprehensive doctrine as either true or reasonable, depending on what that doctrine allows." (Rawls, 1995, 143). Such a statement implies that agents in a well-ordered society are to be understood to be capable of grasping at once both arguments with a neutral basis, and those that are associated with comprehensive doctrine. In order for agents to harmonise public reasons, those associated with arguments for regulative principles of justice, with comprehensive

Consider the way in which Cohen and Murphy, each in his respective way, dismisses the revision in Rawls' later work. It is possible to read Cohen (1997, 17) as setting up a dilemma for Rawls. In a section (section IV) devoted to pointing up ambiguities in Rawls' texts as regards the reach of his principles of justice, Cohen quotes him as saying that Rawlsian citizens "strive to apply" the principles of justice in their daily lives because they,

have a desire to express their nature as free and equal moral persons...When all strive to comply with these principles and each succeeds, then individually and collectively their nature as moral persons is most fully realized, and with it their individual and collective good. (Rawls, 1973/1971, 528, cited in Cohen, 1997, 17).

Cohen on behalf of Rawls responds to this by pointing out that his later work, as I have noted, seeks to shed the idea that there is a common fundamental moral conception that all can appeal to in support of their public and private practices. Rawls can, then, simply disavow his earlier suggestion (captured in the quote used

doctrine it must be the case that a standpoint is afforded them that allows a single evaluative perspective, one that dissolves the divide I have referred to in the text between comprehensive and neutral thinking. I think, however, that this does not save Cohen from the dilemma that, I have argued, he faces. The harmonisation of political principle and comprehensive doctrine that Rawls refers to comes after a neutral principle has been established. It might be that Rawlsian agents will come to see the fit between Cohen's principle of self-restraint and their personal convictions. However, what (at least reasons-based) Cohen wants is a logical implication from a commitment to the difference principle to a personal commitment to self-restraint in the face of incentive payments. He cannot have this, I would argue, because such a logical implication crosses a divide between neutral and comprehensive thinking that, in the derivation of, political principles that, as I read Rawls, ought to be respected.

by Cohen) that he sees principles of justice as both shaping institutional arrangements and informing private deliberations. But a simple disavowal will not mend things, says Cohen. Denying people a common source of inspiration for their everyday decisions leaves individuals, so to speak, morally stranded, the practices of each distinct group being informed by their own, distinct fundamental normative commitments. Cohen says: "it means that the ideals of dignity, fraternity, and full realization of people's moral natures can no longer be said to be delivered by Rawlsian justice." (Cohen, 1997, 17).

Hence, Rawls must, on Cohen's account, supply a bridge between principles at institutional and personal levels if he is to secure a basis for personal conduct that has it cohere with the aims implicit in the specification of institutional rules. So, either: (1) he provides such a common moral basis for all, from which they can work up normative requirements into their everyday lives; or (2) he does away with this common basis and so (because he prevents their being a common moral culture) isolates social groups from one another. Option 1 has Rawls yield to Cohen, while option 2 makes a general social ethos impossible - so preventing Rawls from forging the stability he sees as so essential to his theory. So Rawls cannot, Cohen thinks, escape his monist critique by privileging comprehensive doctrines. Doing so begs the question of what the basis for communal, society-wide mores is. Cohen, as far I am aware, doesn't pursue this line of thought.

However, he is too brief on this aspect of Rawls' work. Once we emphasise, with Rawls, the importance of accommodating disparate comprehensive doctrines in a single polity, we see that we need to confront rather than dismiss this feature of Rawls' position. Once we do that the conflict between principles that do and do

not emanate from comprehensive doctrines, and, respectively, can and cannot be taken to inform personal deliberation, comes into view. And with it an argument against Cohen, even, as I have suggested, on the reasons-based reading.

In section 4.2 I draw on Liam Murphy's criticism of Rawls in order to emphasise the social embeddedness of market conduct, and so complete the second aspect of my argument against reasons-based Cohen. It is worth noting first that Murphy too pays insufficient attention to the divide most emphasised in Rawls' later work between neutral and doctrinal principles.

Murphy is similarly brief in his consideration of Rawls' adjustments to his project. He notes (1998, 255) that Rawls defines the political conception of justice in *Political Liberalism* along dualistic lines: "...it is a moral conception worked out for a specific kind of subject, namely for political, social and economic institutions" (Rawls, 1993, 11). And also notes that "The language of dualism also figures in the account of a political conception of justice as one that stands free of, and is compatible with, a range of "comprehensive views"" (Murphy refers us to Rawls, 1993, 12-15). Murphy goes on to say, however (Murphy, 1998, 255), that,

this central issue for *Political Liberalism* of how a conception of justice can be justified and stable in a society made up of people with a range of different moral, philosophical, and religious views in fact seems to be neutral between dualism and monism.

He refers us to pages 223-27 of *Political Liberalism*, and finds here evidence for his opinion that individual groups can, consistently with Rawls' theory, argue successfully for principles to shape the political arena that shape also their personal lives. Given this it might possibly be the case that an "overlapping consensus" is established whose elements cross political and personal boundaries

for some or, indeed, all groups within the Rawlsian society. There is evidence for Murphy's point here. The search for substantive political principles requires there to be a public understanding of what constitutes acceptable modes of argumentation. Not only must we (presumably us and Rawls qua theoreticians) seek common principles, but must do so in a way that appeals to a universally acceptable conception of criticality ("..each of us must have, and be ready to explain, a criterion of what principles and guidelines we think other citizens...may reasonably be expected to endorse along with us..." (Rawls, 1993, 226)). Thus, in arriving at a consensus some political conceptions will (no doubt) fail to get a hearing due to their being unsupported by reasonable criteria. Those that are successful may, it seems safe to conclude with Murphy, consist of monistic and dualistic elements. So Murphy dispels the idea that Rawls' concern to secure stability in the face of radical moral and philosophical difference is relevant to his (Murphy's) thesis.

Murphy appears to think that Rawls makes room for consensus, while assuming self-contained doctrinal positions, by differentiating between two types of substantive commitment: commitments that reflect established theories of the good; and those that reflect a neutral commitment to democratic mechanisms. One plausible way to combine the existence of substantive political commitments with a belief in the soundness of mastergoal monism (taking that to be Murphy's favoured position) is to take as mastergoal just the political aspect of fundamental convictions. The distributional aspect of the mastergoal, embodying the overriding aims of a society's institutions, on this account is part of the substantive commitments individuals have for their society irrespective of their own distinct

creeds. Monism, in this picture, does not exclude the politically inspired (non-doctrinal) mastergoal from entailing requirements on everyday life. This combination leaves individuals free to have their contending moral, philosophical, and religious visions of the good inform their personal lives, while political principles inform not only aspects of the personal sphere but also a person's vision of the good society, as democratic and egalitarian, along Rawlsian lines. For example, it is possible to imagine an individual whose market behaviour is influenced by a desire to be consistent in his acceptance and application of the difference principle (this reflecting a commitment to the mastergoal); and who buys only products not proscribed by his religion (this reflecting his comprehensive moral vision). This leaves his comprehensive doctrine intact, and has it inform his life appropriately, while political principle infiltrates his life to an extent that does not have it challenge moral, philosophical, and religious convictions. Pogge takes Murphy to task on this. For him, Murphy misreads Rawls because mastergoals are a function of comprehensive doctrines, and political institutions are acceptable only insofar as they further them, and their distinct aims. There is no political or other mastergoal standing over and above those of disparate social groups.

Devotees of diverse mastergoals may be able to agree on the same public criterion of justice, accepting from their separate standpoints that it advances well enough what they believe ultimately matters...To facilitate such a consensus, Rawls himself does not..take a stand on whether there is a mastergoal and, if so, what it is. (Pogge, 2000, 158).

And Pogge's footnote to this reads (n. 38),

His (Rawls') abstinence here instantiates a general effort to keep his theory compatible, as far as possible, with a plurality of more (sic.) comprehensive doctrines by avoiding denial of any views they set forth as well as expressions of skepticism and indifference toward them.

Hence, Pogge denies (what Murphy, as mastergoal theorist, requires) that a Rawlsian society has a mastergoal at all. And this seems right. The presumption builds into Rawls' theory a dangerous understanding that there is a primary theory of the good, to which other theories are underlings. It is precisely this danger that Rawls wants to avoid in his rejection of utilitarian and perfectionist theories of justice.⁴²

According to Pogge, then, the mastergoal monist requires a common understanding of what promotes the just society, mastergoals are a function of comprehensive doctrines, so the fact of plurality precludes their being a single society-wide mastergoal. There just is not then a master goal available for people to feed in to their personal lives. Murphy, however, need not feel thwarted by Pogge's remarks. He can reply that the type of mastergoal he refers to need not be taken to emanate from a comprehensive doctrine. The mastergoal he requires need only relate to political convictions, on the understanding that democratic and society-wide distributional aims have a basis, not in moral, but in a political conception. This seems to put Pogge and Murphy in deadlock, and I'm not sure

⁴² It should be said that in his response to an "expensive tastes" objection to his theory Rawls talks about people having responsibility for their tastes. They ought, he says, to adjust their expectations and exercise restraint accordingly given facts about the scarcity of resources and others' legitimate claims. However, this type of adjustment does not require a change to a personal (or group-inspired) view of what constitutes the good on the grounds that it conflicts with a higher, substantive view of the good life.

how to move on from this. The discussion does, however, indicate that Murphy may well be a little hasty in his rejection of the importance to his argument of the limited place of comprehensive doctrines in Rawls' later work.

In the next section I present the second aspect of my argument against the reasons-based version of Cohen's argument. I draw on Liam Murphy's defence of Monism in order to emphasise the social embeddedness of personal market choice.

§4.2 Second aspect of the argument: the embeddedness of market-based decision

In section 4.1 I argued that due attention to Rawls' aims in his later theory reveals a division between the personal and the political that hinges on the understanding that principles for personal conduct stem from fundamental convictions; and that this basis makes them unsuitable for institutional design. An accommodation of the "fact of pluralism" and stability is not achieved by so grounding political principles. This creates problems for Cohen even on the reasons-based account, according to which he requires Rawlsian citizens to address reasons for action. These considerations have been expressed in a way that emphasises the neutral basis of institutional principles. My focus in this section is on the social embeddedness of market choice, and how an emphasis on its embeddedness supports the view that we ought to recognise a difference between the bases and practical status of institutional and personal principles. I draw again on an aspect of Liam Murphy's defence of his thesis.

In support of his position, Murphy responds first to the argument in favour of dualism that Rawls offers in lecture VII of Political Liberalism, "The Basic

Structure as Subject." There, Rawls makes the following points. The first two are not considered at all by Murphy who opts to move straight to (what is here) the third point - that concerning "a division of labour". The first two are, however, important, as will be seen. Noting that utilitarianism is a theory that applies a common principle in the assessment of political institutions, economic arrangements, and individual acts alike Rawls says that "The first principles of justice are plainly not suitable for a general theory," (Rawls, 1993, 261). And the difference is that justice as fairness is informationally sensitive to intrinsic differences among distinct parts of society. Rawls says,

These principles require...that the basic structure establish certain equal basic liberties for all and make sure that social and economic inequalities work for the greatest benefit of the least advantaged against a background of fair equality of opportunity. In many if not most cases these principles give unreasonable directives. (Rawls, 1993, 261).

He expands by pointing to the way in which associations – he gives churches and universities as examples – differ in terms of their aims, and how this makes the application of common principles to their internal organisation inappropriate. While societal relations among associations are appropriately regulated by a single set of principles, taking these to govern their internal operations fails to recognise the distinct way in which their individual goals shape their objectives. What is to be recognised as a claim against another, for example, within an association will be legitimately determined by, say, what is taken to be its *raison d'être*. But more than this "it seems natural to suppose that the distinctive character and autonomy of the various elements of society requires that, within some sphere, they act from their own principles designed to suit their peculiar

nature." (Rawls, 1993, 262). Hence, Rawls does not only defend the distinction between principles of justice and those that regulate associations on instrumental grounds. There is a relationship between the "character" and the "nature" of distinct associations and the principles that govern them. Combined with a respect for their "autonomy" the picture is that the restriction of principles of justice to the basic structure provides the potential for an accommodation of associations whose *modus operandi* reflect (often) deeply rooted convictions about their purpose, significance and role.

The next reason Rawls gives for the distinction of principles is the following (Rawls, 1993, 266). Considered just as such, a set of transactions or, more generally, individual actions may be, taken singly, just according to a particular criterion. This does not prevent the outcome generated by the actions in combination producing an unjust circumstance. Declining relative income for one concentrated group of workers – unacceptable according to a general principle – may well be consistent with each individual action, in the series of actions leading to the decline, being permissible by some principle whose application is to individual choice. Given this possibility it is necessary that we have principles not only for individual behaviour but principles that will enable us to comment simultaneously on what is happening globally and relatively. This prepares the way for, but is distinct from, the third defence - that associated with what Murphy calls the "division of labour". Rawls says, "That there are no feasible rules that it is practicable to require economic agents to follow in their day-to-day transactions that can prevent...undesirable consequences. These consequences are often so far in the future, or so indirect, that the attempt to forestall them by restrictive rules

that apply to individuals would be an excessive if not an impossible burden." (Rawls, 1993, 266). Hence, the fact that local interactions are not transparent in their consequences makes it unfeasible to require individuals to seek out the information necessary to make a global assessment of their action. The application of general principle to activity at the individual level would require such information seeking and so would be impossibly burdensome. But this is, I suggest, a secondary consideration. That what happens at individual or local levels, and their aggregated consequences may come apart in terms of their justness is the primary concern. That a division of labour is appropriate is a consequence of this prior fact. That individual transactions could lead to unjust global outcomes is, of course, insufficient grounds to worry about global consequences. But coupled with the view that unregulated interaction will, however, lead to "oligopolistic configuration of accumulations that succeeds in maintaining unjustified inequalities", we not only have an argument for general principles to regulate background conditions, but, prior to that, a justification for justice-preserving background institutions, to which principles of justice can be applied. "(W)e require special institutions to preserve background justice, and a special conception of justice to define how these institutions are to be set up." (Rawls, 1993, 267).

Murphy, emphasising the division of labour aspect of Rawls' discussion here, (Murphy, 1998, 259), is unconvinced that what Rawls has to say helps to address the monist-dualist dispute: "But none of this supports the idea that different normative principles apply to institutional design on the one hand and personal conduct on the other." In support of his view Murphy discusses a

"conception of justice that rejects dualism, but embraces the ideal of the division of labour." (p. 262). Murphy employs a principle of "weighted beneficence" a principle that attaches more weight to the worse off in comparisons of distributions. The principle is utilitarian in that according to it it ought to be the case that weighted aggregate well-being be promoted, and it is monistic because this weighted aggregation procedure (by assumption) is applied to both institutions and to individual choices. Nonetheless, says Murphy, it will "make eminent sense" (p. 263) to concentrate responsibility for achieving desired outcomes at an administrative level - responsible for setting taxes and co-ordinating transfers - while relieving individuals of much of the burden of steering the economy towards just allocations. And this seems correct. We cannot argue for a separation of principles, between institutions and every day choices, on the grounds that it is more efficient to free up individuals from the burdensome task of active moral stewardship. But this, while no doubt a knock-down argument for some variants of dualism, does not respond to aspects of Rawls' opinion (as expressed in "The Basic Structure as Subject") that are important for his theory. Rawls does present instrumental arguments for a separation of principles, in a division of labour argument that Murphy here convincingly overturns. But, as noted, Rawls does not rely on the division of labour point, it is, rather, merely an upshot of his more basic points.

Note that part of Lecture VII (Rawls, 1996, 262-265) is a response to the libertarian rejection of systematic co-ordination of economic activity at the institutional level for the purposes of distributive justice, and the notion that "special first principles are required for the basic structure." (p. 262). The

libertarian argument, as Rawls presents it , has, as a primitive feature, the assumption that a single set of rules can be applied to all transactions in order to assess whether the outcomes they generate are just. A basic datum in the theory is a conception of transactions as a single type of item to which, as such, a single set of rules can be applied. For Rawls this is an inappropriate starting place. Basic to his theory are not transactions as such, but transactions embedded in rule-laden contexts. On page 267 of *Political Liberalism* Rawls says that,

..fair background conditions may exist at one time and be gradually undermined even though no one acts unfairly when their conduct is judged by the rules that apply to transactions within the appropriately circumscribed local situation.

This can be interpreted as a reference to local circumstances given a background assumption that general rules at an institutional level already exist. But this is not Rawls' intention. Consider the rest of the paragraph - interspersed with my comments:

The fact that everyone with reason believes that they are acting fairly and scrupulously honouring the norms governing agreements is not sufficient to preserve background justice. (ibid.).

This directly contradicts what the libertarian would say, with "fairly" given a Nozickian gloss. And if this is a challenge to the libertarian, on his terms, then it would be appropriate for us to understand these local transactions as taking place independently of any administrative hierarchy. Rawls goes on,

This is an important though obvious point: when our social world is pervaded by duplicity and deceit we are tempted to think that law and government are necessary only because of the propensity of individuals to act unfairly. But to the contrary, the tendency is rather for background justice to be eroded even when individuals act fairly: the overall result of separate and independent transactions is away from and not toward background justice...Therefore, we require special institutions to preserve background justice, and a special conception of justice to define how these institutions are to be set up.' (ibid.).

This recommends a reading that sees the local transactions referred to as taking place prior to the setting up of a general institutional framework. The rules applied "to transactions within the...local situation", "the norms governing (local) agreements", are, then, to be taken to be, in this quote, prior to more general rules that apply across local contexts. The point is, and this is the basis of Rawls' offensive here against the libertarian, that economic interactions considered fair from the perspective of the locale are not guaranteed to promote fair outcomes at a general level. This is, empirically, because general structural facts such as inequalities of bargaining power, can (or, we can say, with some conviction, will) create unfair outcomes in the aggregate. For example, someone's job acceptance can, taken by itself, appear to be just, whereas once we take account of the pressures on the individual to accept the work, pressures originating from outside the local context (such as a decline in demand for the output in her region of the industry in which she currently works, due to monopolisation in another region) the transaction may not appear so fair (Rawls gives an alternative labour market example on page 267). But, importantly, this result does not depend on empirical assumptions alone. It is basic to the argument that transactions with distinct locations - where what defines a location are the norms under which transactions occur - cannot be taken to satisfy a higher-level norm, one that breaks down the

distinction between different contexts. Were this to be possible then the libertarian would be let back in. He could say that he recognises local norms but what takes precedence, for theory, are the transactional principles that secure justice at the general level.

It is essential, then, if his response to the libertarian is to work, that Rawls accommodate, rather than override, local principles in his argument for principles of justice. It is, at the same time, imperative that such institutional principles be established because (something the libertarian does not allow within his argument) fair principles of cooperation that seem to operate well locally do not offer protection from injustices that come from background events. The latter is also an important point from the perspective of a discussion about monism and dualism. It might be suggested that once the need for institutional principles is recognised, the local norms can be overridden, supplying a monism that Rawls need not object to once he has convinced us of the need for institutional concern for distributive justice. To see why this is not the case let us consider the following distinctions - distinctions that are not, I think, made explicit in Murphy, nor indeed in Cohen.

The rules that constitute the market as an institution detail the permissions and contractual and other obligations of market participants as they engage in production and exchange. These rules govern everyday individual conduct in the market context. All trades, for example, must be voluntary. These rules are, in turn, influenced by the principles for institutional design ground out by the original position device, and are thus rules all can accept as legitimate restraints on what they can permissibly do. Restrictions on monopoly power, for example,

intended to avoid unequal democratic representation entail restrictions on the conduct of firms they will accept if they accept the underlying rationale. Individuals not directly involved in production will nonetheless be aware of and, similarly, accept the restrictions as legitimate, given a commitment to their underlying rationale. This sort of influence on conduct is general to all market participants, in that all who engage in exchange or production must abide by them. There are other influences, however, that do not apply to all market participants. What I have in mind are influences associated with comprehensive doctrines (in Rawls' sense of the term). For example, choices associated with veganism, religious codes, environmentalism, fashion, and culture are all informed by ethical outlooks. The point is not to contrast the compulsory nature of rules of market exchange with the discretion that applies to personal commitments, and tastes. The distinction is, moreover, crucial if Rawls is to preserve a division between restrictions on personal behaviour that emanate from neutral reasoning (behind-the-veil), and moral restraints whose support is embedded in comprehensive doctrines.

The market context is an instance of Rawls' concern to take as datum doctrinal distinctions across different groups. If he helps himself to the view that general restrictions associated with principles for institutional choice can influence personal behaviour *as part of the* comprehensive doctrine of all individuals then he fails to secure a distinction between restrictions associated with neutral reasoning, and restrictions associated with comprehensive doctrine. That is, general restrictions, such as a requirement to abide by personal tax codes, inform personal choice, but the reason for this cannot be that their rationale is

integrated with the other constitutive elements of a comprehensive doctrine. This would assume the compatibility of the support for the details of institutional design and fundamental moral conviction, precisely what Rawls cannot assume if he is to move away from the instability in his original account. Hence, central to my argument is that normative considerations that feature in ground level market activity are wedded to comprehensive moral outlooks, and that the emphasis Rawls places on accommodating standing comprehensive doctrines carries over to the domain of market choice. This being so there is a divide between what can be acknowledged as a restriction on market outcomes, engineered parametrically, through institutional adjustment, and the pressures that bear on individual behaviour, and have a basis in the fabric of personal moral commitments. Preserving this divide, as Rawls must, imposes dualism as a requirement in his theory, rather than an option he favours. It also delinks his dualism from arguments against it that point to ambiguities in the specified reach of the basic structure.

§5. A “Delimitative” Response

The picture that emerges is of a market participant who is constrained in two ways. First, she is constrained to comply with restrictions on her conduct associated with the specific details of her society’s institutional architecture. She must comply with tax regulations, company law, contractual obligations, and the details of the rules that ensure that market exchange is voluntary, and that

property rights are respected. Rawls presumes a market economy⁴³ and so it is reasonable to suppose that an understanding of and respect for the general restrictions associated with market participation are embedded in the outlooks of Rawlsian citizens. This does not mean that the restrictions associated with policy, the details of which are inspired by the principles of justice, have a rationale that is similarly embedded in comprehensive outlooks. Indeed, it is this latter type of restriction the rationale for which cannot be assumed to be integrated with comprehensive doctrine. To assume so fails to ensure a necessary divide between rationales of neutral standing and those emanating from a comprehensive view. In order to accommodate this contrast, a contrast I am suggesting Rawls must acknowledge, it is essential that there is a related distinction between two types of normative requirement. I will call these “delimitative” and “personal” normative requirements. A delimitative normative requirement need not feature in, or play a role in the formation of personal choice in everyday situations, situations in which

⁴³ Rawls sees his theory as applying in principle to both socialist and capitalist systems, differentiated by the extent of private ownership, hence the market he presumes (at Rawls, 1972/1971, 7) is a system of exchange and production with the degree of private ownership unspecified. Luck egalitarian readings of Rawls, for example Scheffler (2003), tend to identify his use of the market, however, with a response to Conservative thinkers who doubt the ability of egalitarians to adequately respect the personal responsibility the market system, with its devolved responsibility for the management of personal economic affairs, engenders. On this reading the market, in Rawls, rather than being indifferently located within socialist and capitalist polities is understood to be the market as it exists in modern liberal democracies. What I have to say is insensitive, I think, to these two readings of Rawls’ intentions. It is relevant to note here Rawls’ view that it is only by supposing a market system “..that the problem of distribution can be handled as a case of pure procedural justice. Further, we only gain the advantages of efficiency and protect the

a choice or decision is to be made from a first-personal perspective. An example of a delimitative normative requirement is, relevantly for the discussion, a requirement to respect the negative rights (rights to non-interference) of others, with respect to their personal property. While (widespread) obedience to such a requirement is necessary to the long-term stability of a system of market exchange, it is not a requirement that a person need ever bring into play in his or her actual market choices. The requirement sets parameters to what can permissibly be done, restricts the sort of choices that can, permissibly, be made, but it need not feature in a choice, or a piece of practical reasoning in the market context, that is itself consistent with the parametric constraints the delimitative requirement sets. Requirements associated with the policies, laws, and regulations that are intended to give material shape to Rawls' two principles of justice are delimitative in this sense.

The distinguishing feature of a *personal* normative requirement, as I intend it, is that it features in, plays a performative role in, personal, practical reasoning. Rather than placing limits on what might be permissibly chosen, a personal requirement shapes, or contributes to the process of actual evaluation, at point-of-choice, as a constitutive aspect of the process of deliberation itself. These distinctions are helpful in characterising the form a person's practical reasoning should have if it is to satisfy, for the market context, the requirement that there be a distinction between requirements that have their basis in comprehensive vision, and those that find their support in neutral, behind-the-veil, argumentation.

important liberty of free choice of occupation.” (1972/1971, 274). None of these properties rely, in principle, on private ownership of productive means.

Personal requirements will include those associated with moral principles and commitments that inform what a person buys, where she is prepared to invest, and so on. These sorts of requirements associated with comprehensive doctrine can then have a different pedigree to those, delimitative, requirements that, in the economist's parlance, restrict her feasible set. This sort of distinction is a must, I would argue, if Rawls is to ensure that comprehensive doctrines do not infiltrate the original position, and principles so derived do not automatically harmonise with comprehensive doctrines.

I want to consider now the implications of the distinction between delimitative and personal normative principles for the reasons-based version of Cohen's position⁴⁴. Is Cohen to say that the moral requirements he thinks ought to be in play in the market place are "personal" requirements, in the sense in play here? It seems that he must do so if they are to be thought of as pressing for market actors in the personal domain, in their everyday lives. If he does say this, however, he is characterising these requirements as emanating from a person's comprehensive doctrine. If he denies this then the principle must emanate, rather, from the reasoning associated with the neutral vantage point of the original position, and if he claims this then he threatens to break down the distinction, so crucial in Rawls, between the normative outlook connected with a person's comprehensive outlook and principles whose origin is neutral, from outside of that outlook.

⁴⁴ The argument of this section (§5) relates to reasons-based version of Cohen's view, so that references to Cohen are to reasons-based Cohen.

Cohen might want to claim that the personal moral requirement he supports is delimitative, and that because of this the argument here doesn't touch him. He might offer the seemingly plausible argument that a requirement to forego payments made purely for extra effort need not feature *in* personal choice. The associated principle could be one that rules out - on a reasons-based reading of his view - certain reasons for choice. However, characterising the personal requirements he defends as delimitative will not help his case. What he needs to be convincing against Rawls is a requirement that features *in* the personal sphere, in the everyday orientation of Rawlsian market agents. A delimitative requirement does not do this. It is, indeed, on all fours with requirements associated with the policies, laws etc., that are not personal but institutional.

§6. Conclusion

Cohen's general claim is that if a person defends principles of justice for institutions then they must also recognise implicit principles for personal conduct. The way he argues for this in relation to Rawls is to point out an inconsistency between a commitment to the difference principle and conduct that generates unnecessary inequalities. In particular, market participants ought to be prepared to forego material rewards for effort that could be made in their absence. Prominent defences of Rawls on this score, associated with Thomas Pogge and Andrew Williams, rely, I have argued, on non-compulsory readings of Cohen's view. Cohen can, I think, side-step their remarks. He can place a reasons-for-action emphasis on his claims. On this reading the Rawlsian market agent ought to modify her beliefs about what constitutes acceptable reasons for action, and then

outlaw those reasons as a basis for market exchange. This sort of exercise in self-restraint does not require agents to make counterfactual judgements about the possible consequences of their actions, as in the Pogge-Williams view. Indeed, a world in which pure incentive payments are routinely paid and routinely accepted can be one from which the offending reasons have been expunged.

This would seem, at first pass, to leave Cohen's position poorly defended. If the existence of inequalities that he thinks ought to be outlawed by the difference principle is consistent with universal acceptance by market participants of his restriction on reasons, then his ambition of making Rawls more egalitarian is not, it would seem, achieved. But it is perfectly compatible with Cohen's monist thesis for him to see the removal of such inequalities, those associated with pure incentive payments, as an institutional task, rather than a task for the individual.

There is, however, an alternative way to criticise Cohen, one not so easily side-stepped by a reasons-based reading. The embeddedness of market-based decision is important for Rawls, I have argued, not only because, being everyday action, it is informed by agents' comprehensive doctrines. If Rawls is to avoid dissolving the divide between principles supportable from the neutral stance of the original position, and those whose support is embedded in standing, "thick" doctrine, he must distinguish between those principles in terms of their support. In order to avoid dissolving this distinction, I have argued, Rawls must have in place a further distinction between what I have called delimitative and personal normative requirements. The problem this causes for Cohen even on a reasons-based reading can be stated as follows: Either: (1) he sees the personal principle he wants to defend as based in comprehensive doctrine and so of a different pedigree to the

(institutional) difference principle; or (2) he sees the personal principle as having the same pedigree as the difference principle and so imposing a requirement without purchase in the everyday life of the market actor. This, however, can be circumvented, or so Cohen might argue. A delimitative moral restraint places out of bounds prohibited considerations prior to actual deliberation taking place. Cohen could argue that market agents can, whatever their cultural background, place such prior restrictions on what are permissible choice-guiding reasons. The question of the purchase of such restrictions need not arise, then, once market actors are engaged in market exchange. Enlightened Rawlsian market agents will, having seen the truth of Cohen's argument, see as off-limits the reasons he has convinced them ought not to be taken into consideration. But this attempted solution is problematic. What Cohen wants (must) argue for is personal moral restraint. A delimitative moral principle is not one that enters into the considerations of an agent as she is engaged in (in this case market) activity. It sets boundaries around what is permissible anterior to deliberation towards choice, but is not in play in such deliberation itself. So this strategy does not introduce personal moral restraint into the lives of market actors, it merely precludes the possibility of the offending reasons being part of their everyday deliberations at all. It does not address the problem identified as it excludes from the personal domain the principles it seeks to argue ought to be felt there.

What this discussion of Cohen serves to emphasise is the importance in Rawls of the distinction between the institutional context of market choice and the everyday orientation of the market actor. Choice in the latter, everyday setting is informed by comprehensive doctrine. Moreover, what ensures that the distinction

matters is the discretion agents are allowed in the market context. This makes the market a powerful way to achieve both the preservation of a thoroughly private space, informed by comprehensive doctrine, and an institutional framework whose permissions and prohibitions rely on extra-doctrinal considerations. Cohen's thesis threatens to break down this divide and so undermine what is paramount in Rawls, that (reasonable) comprehensive doctrine inform personal evaluation, in the personal sphere, while the design of basic institutions has non-doctrinal foundations. Discretion in the marketplace is central to the way in which Rawls preserves this distinction in the market context, and such discretion is central to the next chapter.

Chapter 3

Market Participation and Embedded Critique in Rawls

Introduction

A stable liberal democratic society is not one from which the potential for dissent has been expelled. Those who reluctantly accept the institutional status quo do not exercise the concerns that motivate their disapproval, they, rather, exercise restraint, they moderate the practical force of the concerns that would otherwise inform their conduct. However, Rawls' conception of the impartial standpoint a person must adopt in order to engage in public debate fails to accommodate the practical status of the concerns that motivate dissent. Imaginatively shorn of the distinctive outlook that forms the basis of those concerns the person, in deliberative mode, sets their practical significance aside. Those concerns cannot be addressed, then, from within the outlook that supports them, but must, if they are to be overridden, be worked on from outside of the person's outlook. Suppression is the only option for one who accepts that those concerns ought not to be allowed unhindered practical force in everyday conduct, and this removes what ought to be present in the ongoing stance of persons in a stable liberal democratic society: the potential for dissent.

In this chapter I examine the extent to which this weakness, in Rawls' representation of the problem of stability, can be overcome by appealing to the

properties of a market system, and the associated characteristics of a market actor. In section one I consider the place of the market in Rawls, emphasising the role the market plays in securing a personal domain in which choice is informed by conceptions of the good. My intention here is not merely to repeat, with further textual evidence, the point that market choice is informed by conceptions of the good. What I want to emphasise is the necessity to Rawls of the market, that it is not something that can be easily removed from his thesis without doing some structural damage. This helps to show the worth of pursuing a conception of embedded critique that focuses on the market. Section two establishes the character of what I call *reason-based autonomy* in the context of the market, and works out its implications for a conception of reflective critique. An aspect of Ronald Dworkin's resource egalitarianism is drawn on here in order to bring out the meaning of reason-based autonomy. In section three the argument draws on a "metapreference" approach to preference formation. In section four I attempt to generalise to other, non-market, contexts of choice, the conception of embedded critique I defend for the market setting, while section five considers how the argument against Cohen, presented in chapter three, might seem to need adjustment in the light of what is said in this chapter. Section six concludes.

§1. The Market in Rawls

A central aspect of Rawls' constructivist method⁴⁵, and what helps to distinguish his approach, quite clearly, from Immanuel Kant's (Cf. Rawls 1989), is the

⁴⁵ His constructivism is laid out most explicitly, perhaps, in Rawls (1980, 1989). The general character of Rawls' constructivism is outlined in an illuminating way in Cohen (2003, §c).

inclusion at a fundamental level in his theory of presumed facts about the society to which the principles of justice are intended to apply. One such presumed fact is that economic activity is organised through competitive markets. Indeed, Rawls confirms early in a Theory of Justice that he sees a system of competitive markets as an example of a major social institution (1972/1972, 7):

By major institutions I understand the political constitution and the principle economic and social arrangements. Thus...competitive markets, private property in the means of production, and the monogamous family are examples of major social institutions.

Later he tells us why he includes a market system among the basic institutions to which his principles are to apply. It is only by supposing a market system, he says,

that the problem of distribution can be handled as a case of pure procedural justice. Further, we...gain the advantages of efficiency and protect the important liberty of free choice of occupation. (1972/1971, 274).

So, rather than simply acknowledging competitive markets as a datum he is forced to take as empirical input, Rawls opts to include the market as a major social institution because of what he sees as its desirable properties. First, a system of voluntary exchange that permits ground level transactions to be based on personal preferences affords a distinction between aggregate outcomes, background, market-oriented policies, and personal decisions. So long as the way in which personal decisions generate aggregate outcomes is consistent with background rules that are considered just, then the outcomes themselves will be just. This, the procedural character of the market, is not, however, only reflected in the distinction between uncoordinated choice at ground level and justice preserving

background rules. I will return to this point shortly. Second, according to the above quote, Rawls views efficiency as a “fundamental social problem” (1972/1971, 6) that a realistic set of principles must seek to accommodate, and he accepts, it seems, the mainstream economic results that competitive markets secure efficient allocations of goods and productive inputs.⁴⁶ Regarding Rawls’ third point, in the above quotation, that markets “protect the important liberty of free choice of occupation”, I note only that this view is much debated for capitalist economic systems.⁴⁷

Returning to the first point, that the market lends itself to a procedural interpretation of what social and economic justice requires, there is an alternative

⁴⁶ That he accepts arguments for market efficiency is suggested by the view he expresses in the quote in the text (1971/1972, 274). He also says, however, that, “..if the received doctrine is..mistaken..I hope that for the purposes of the theory of justice no harm is done.” (p. 234) We can, I think, reconcile these two remarks by viewing Rawls’ acceptance of economic theory as tentative; or, perhaps better, as tentatively held by “the citizen who is trying to organize his judgements concerning the justice of economic institutions.” (ibid.). The relevant area of economic theory is available in most microeconomic principles texts. Eg. Cowell (2006, 238, Theorem 9.4)) and Varian (1992, Ch. 17), incorporates the classic results in Debreu (1972/1959) and ensuing modifications. Cowell’s theorem 9.4 states the case for efficiency in a way that makes explicit the need to exclude “externalities” (third-party effects of economic activity) and “non-private goods”.

⁴⁷ Reiman (1987) is a good source for the thoughts of Roemer and Cohen on this issue. I note also that Rawls intends his theory to apply indifferently to polities with regard to the extent of state ownership of the means of production. The contrast he draws in the introduction to the revised edition of "A Theory of Justice" (1999) is between a socialist liberal society (state-centric ownership) and a property owning democracy (dispersed ownership).

way of making the point. In a way commented on by David Miller (1990, chapter 3), a significant feature of a system of competitive markets is that it is neutral as to which conceptions of the good individuals apply in their market decisions⁴⁸. This facilitates a division between market-oriented public rules designed to secure just outcomes and action inspired by comprehensive doctrines at ground level, and this, I would argue, is an under-emphasised way in which the market is helpful to Rawls. The nature of the contrast can be explored by way of comment on what Kymlicka (1990) sees as Rawls' incipient luck egalitarianism. According to Kymlicka, Rawls is motivated to present a theory of justice that is sensitive to personal voluntary choices, while being blind to differences in unchosen (personal and social) circumstances. As Scheffler puts it,

Will Kymlicka offers a clear description of this putative progression of (luck egalitarian) ideas. Kymlicka says that "[o]ne of Rawls's central intuitions . . . concerns the distinction between choices and circumstances," and that Rawls is "motivated" by the desire to produce a theory that is "ambition sensitive" but "endowment insensitive"—a theory that makes people's fortunes depend on their choices but not on their natural endowments or other unchosen circumstances... (Scheffler, 2003, 8. Page references to Kymlicka, in the original, have been omitted).

As Scheffler also notes, Kymlicka presents, in fact, two theses:

1. Rawls is an originator of a luck egalitarian approach. Strong evidence for this is his rejection of a system of "natural liberty" because the associated unbridled free markets will fail to compensate individuals for inequalities that arise as a result of

⁴⁸ Miller argues that markets are not sufficient for neutrality, but I needn't address this in

unchosen factors that, being unchosen, place arbitrary limits on persons' opportunities.

2. Rawls, perhaps because he is unaware of his role as originator of a luck egalitarian stance, fails to adequately separate the influences of chosen and unchosen factors in persons' lives. A prime example of this is his use of primary goods as an index of inequality. The allocation of primary goods is directed by the difference principle's concern for relative equality, without reference, for example, to the indolence of couch potatoes, or, say, the special needs of the disabled,⁴⁹ so, using Dworkin's way of putting things (Dworkin (2000)), the option/brute luck divide is poorly drawn.

Scheffler goes on to doubt whether Kymlicka has got Rawls right in this regard. He would tend to agree, he says, with Elizabeth Anderson (1999) that luck egalitarianism requires a higher-order moral perspective that embraces the distinction between choice and luck. An approach that takes the latter distinction as a primary building block for egalitarian theory cannot take into account the relevance of the specific context in which a choice/luck distinction might be in play – and this, *inter alia*, points up the need for the luck egalitarian to provide a prior moral basis for a focus on choice. Scheffler is convincing here⁵⁰ and I think

order to make my points.

⁴⁹ This line of criticism, as is well known, is pursued by Sen (most recently in Sen (2000)).

⁵⁰ A central quote is, Scheffler (2003, 33): “The intuitive view, I believe, is that the fairness or unfairness of differences in advantage resulting from, on the one hand, factors

that we should see that in Rawls the choice/circumstance distinction is very much contextualised; and what provides that context, I would argue, is both his concern to divide neutral and non-neutral orientations in personal choice, and the way the market helps him to secure that distinction in everyday life. Important, if this point of view is to be seen as in any way correct, is the idea that when a person is engaged in personal choice in the marketplace, what informs her choice is her personal outlook. I would argue that this is the case, in Rawls, for two reasons:

1. If conceptions of the good are to have the protected status they are supposed to have in Rawls then it seems inappropriate to see neutral considerations as invading the personal sphere when it comes to market deliberation. One way to incorporate reasons-for-choice that have a neutral basis in market choice is to integrate them somehow with reasons that emanate from comprehensive doctrine. This, however, would threaten to break down the divide that ensures comprehensive doctrines do not seep into the foundations of the neutral outlook;

beyond people's control and, on the other hand, people's voluntary choices, is highly dependent on the prevailing social context and institutional setting. For example...most (would not) think it unfair if a naturally gifted professional athlete were offered a more lucrative contract than his less talented teammate. On the other hand, most people would consider it outrageous if an emergency room doctor left an injured patient untreated simply because the patient's injury resulted from a foolish but voluntary decision. And few would think it acceptable to deny legal counsel to an indigent defendant on the ground that her inability to pay for an attorney was the result of poor financial decision making on her part. If this is correct, then the intuitive credentials of luck egalitarianism cannot be taken for granted, and it is all the more important for luck egalitarians to explain how exactly their position is supposed to follow from a conception of persons as having equal moral worth."

2. I argued in chapter three that one, important, way in which Rawls argues against a Nozickian position is his presumption that market transactions are culturally embedded. This stymies libertarian attempts to construct a principle of justice that applies to a theoretically primitive conception of transactions just as such.

Hence, the usefulness of the market, for Rawls, is, arguably, that it helps preserve a personal domain in which any ethical considerations that are in play emanate from standing comprehensive doctrine. Neutral ethical evaluation applies not to personal choices but to overall outcomes. Further, where overall outcomes are deemed unacceptable from the point of view of justice any adjustments that are to be made must be to the background “rules of the game”⁵¹ and not at the level of personal choice. This requirement is reinforced by a conception of a system of market exchange and production that emphasises the privileged status of personal values, and personal beliefs at ground level.

The central thing to retain, from these brief remarks about the role of the market in Rawls, that will be useful as we proceed, is the way in which personal market choice is informed by comprehensive doctrine. The agent is non-neutral

⁵¹ The term “rules of the game” to describe the basic conditions for market interaction is commonly used, for example Friedman 1970: “there is only one social responsibility of business – to use its resources and engage in activities designed to increase its profits so long as it stays within the rules of the game, which is to say, engages in open and free competition without deception or fraud.”

when engaged in market-based practical deliberation (the thought is expressed less starkly later).

§2. Limited and Unlimited Reason-Based Autonomy in the Context of Market Exchange

I want to begin now to attempt to address the issue of the neutralisation of the motive status of doctrinally based claims, in the context of a Rawlsian society, by presenting a particular characterisation of the market participant. It will be helpful, first, to describe what I take to be the relationship between, what I call, limited and unlimited reason-based discretion in the market setting. Note, first, the place of a person's *reasons for making a choice* (what might be called "the because" of her choice)⁵² in the process of preference formation.

§2.1 Reasons, Evaluative Criteria, and Preference

When a person makes a choice in a market context evaluative criteria are brought into play. In a decision about which of two houses to buy, for example, the houses may be compared according to the number of rooms, proximity to transport links, and so on. Each possible comparator is an evaluative criterion that, generally, informs the way a person compares two objects. There is a relationship between the evaluative criteria a person brings to bear on a choice and the reason(s) she has for making the choice. A person about to start a family will have more

⁵² Someone might say, "I am choosing a new car *because* my old one no longer works". That their old car no longer works is here the reason for the choice. Alternatively a person might say, "I am choosing a new car because I have money to burn". Theirs would be an alternative reason for making the choice.

concern, perhaps, for the safety of a neighbourhood, or the size of the rooms, when choosing a new place to live than someone looking to rent a house out to students. The formation of a preference in the marketplace, where by preference I understand an ordering of possible options (an ordering that will, or would, inform actual choice, were the occasion to arise), is informed by these evaluative criteria. In turn, which personal evaluative criteria that are in play in the process of preference formation is influenced by the reason(s) the person has for making the choice (the reason she has for ranking those possible options rather than focusing her attention, at that time, on other objects of choice). In any given choice scenario, then, a person arrives at a preference, a ranking of available options, by bringing into play evaluative criteria that reflect her underlying reasons for orienting herself in that way. This in no way attempts to elide the complexity, often noted, of the way that evaluative criteria are brought to bear in a choice situation. Nor does it have to avoid the issue of incommensurability (dealt with in an interesting way in Raz (1999, Ch. 3)). All that is needed here is to think of what a person would ever are the evaluative criteria she is applying. How they operate *in the act of* comparison is not a relevant issue.

Having located the place of reasons-for-choice in the process of preference formation, I want in the next two sections, to draw out the relationship between, what I call, limited and unlimited reason-based discretion in the context of market exchange. As a first step, we can get a good idea of what is meant by reason-based discretion by looking at how it features in Ronald Dworkin's version of resource egalitarianism.

§2.2 Reason-Based Discretion in Dworkin's Resource Egalitarianism

An important ingredient in Dworkin's project is his attempt to show how, through a characterisation of the properties of a market system, a normative political theory can be defended that accommodates, in a non-antagonistic way, both liberty and equality. A way to see how Dworkin achieves this is through an interpretation of his thoughts on expensive tastes. Under conditions of scarcity a person who cultivates expensive tastes will make a greater call on resources than others whose desires are more easily satisfied. If we imagine that a stock of all-purpose resources are centrally held for distribution according to the principles a society accepts, then the decision will be, in a simple welfare-egalitarianism, that some centrally held resources ought to be released for the person with the expensive tastes, in order to maintain equality of welfare. This, given scarcity, makes those resources unavailable to others, thereby imposing a cost on them. A pertinent question to ask is how it is decided how much additional resources the person with expensive tastes ought to receive. If it is to be based on the person's subjective assessment, of the shortfall between the resources needed to bring him up to a welfare level equal to others and the resources he has, then there is no limit in principle to the resource transfer, and associated cost that others must bear. Moreover, a welfare egalitarianism that attempts to fix this by introducing an objective assessment, based on, say, a theory about the relationship between higher gustatory tastes and personal welfare, will not bring into play, in the calculation, the preferences of other people. In this way, equal concern (the "sovereign virtue") is not shown to all. The desires of the generality are muted in the determination of the resource needs of the bon vivant.

Contrast the way the market makes this kind of decision. Here expensive tastes are registered in the marketplace as claims on resources, where the intensity of desire for a particular good is reflected in a preparedness to forego other desired goods, including leisure. Similarly, the desires, for other purposes, of other individuals for the resources required to satisfy expensive tastes, will be registered by claims made upon those resources. Who gets what is decided via competition in which the person with expensive tastes will be able to attract only those resources he is prepared to outbid others for, perhaps taking less leisure in the process in order to bolster his claims. The corollary of a higher price paid by the bon vivant for the resources he needs is the amount of other resources he must forego, these being released for general use (added to the extra output he, presumably, generates due to reduced leisure). In this way he pays the aggregate opportunity cost to all others of the resources he acquires. As Dworkin puts it,

Under equality of welfare, people are meant to decide what sorts of lives they want independently of information relevant to determining how much their choices will reduce or enhance the ability of others to have what they want. That sort of information only becomes relevant at a second, political level at which administrators then gather all the choices made at the first level to see what distribution will give these choices equal success...Under equality of resources, however, people decide what sorts of lives to pursue against a background of information about the actual cost their choices impose on other people and hence on the total stock of resources that may fairly be used by them...(T)he contingent facts of raw material and the distribution of tastes are...background facts about what equality of resources, in these circumstances is...The market character of the auction...is an institutionalised form of the process of discovery and adaptation that is at the centre of the ethics of that ideal...(T)he true measure of the

social resources devoted to the life of one person is fixed by asking how important, in fact, that resource is for others." (Dworkin, 2000, 69-70. Cf. 84-85 & 122).⁵³

It is "sovereign" he says (p. 72) in the argument that "people enter the market on equal terms" (ibid.). Were there to be unequal initial claims, or the amassed wealth of a few gave them greater command over resources *ab initio* then equal voice would not be given to the preferences of market participants. Their desires would have unequal weight in the calculation of social cost, and so the criticism of the welfarist response to expensive tastes, considered above, would apply here also.

It is essential then to Dworkin's argument that conditions be such as to secure participants' preferences not be muted in the market process. One essential condition (that holds only for an ideal market) is that initial claims be equal and the process not reflect growth, over time, in, for example, asymmetric bargaining strengths. In addition, the market shows equal concern for all participants if it is maximally sensitive to people's preferences. Maximal sensitivity to preference helps to determine, Dworkin says, the "true opportunity costs" (p. 150) of people's holdings. And this brings us closer to the place of reason-based discretion in Dworkin's work.

⁵³ On page 72 Dworkin appears to suggest that the use of a central allocation used to mimic the market process, rather than the market process itself, would suffice to deliver equality of resources. This seems to undermine the reading that has him say the market process must be gone through for equality of resources to be achieved. I think, however, that he is considering here what the practical political implications of his insight might be, and countenancing a real-world modification that does not challenge his theoretical view that the process is a *sine qua non* of justice in resource allocation – but see p.112, paragraph 3.

He asks us to imagine a group going to buy land on which to build a football stadium (p. 150). The price paid for sufficient land would likely be less were the land sold as a single lot than were the lot to be divided into smaller portions. This is because the smaller lots will be capable of being put to a larger number of competing uses than the single football stadium sized plot. The sale of smaller lots also seems to serve better the objective of achieving equality of resources because "an auction is fairer...when it offers more discriminating choices and is thus more sensitive to the discrete plans and preferences people in fact have." (pp. 150-151).

The lack of discrimination in the object for sale makes redundant preferences that exist and would, were they to be brought into play, effect a better estimate of the opportunity costs of the resulting resource holdings. Not only ought goods to be identified in ways that makes them maximally divisible but also the uses goods can be put to ought to be as wide as possible. A restriction on what can be done with a good makes redundant a type of preference a person might bring into play in deciding whether to purchase it. If we picture market-based choice as a comparison of options in which the chooser bases a comparison on evaluative criteria, as in the threefold distinction made in section 2.1, above, it seems straightforward that Dworkin would see the unhindered application of self-selected evaluative criteria, on the part of market participants, as a beneficial feature of a market system. It serves the dual purpose of showing equal concern - types of criteria are not, from the outset, excluded - and helps to determine true opportunity costs (thereby achieving an appropriate combination of equality and "liberty/constraint base" (Chapter 3, *passim*)).

Moreover, if we describe the relationship between personal reason(s)-for-choice and evaluative criteria in the way I did in the previous section then it is a short walk to the view that reason-based discretion is a desirable property of a market system for Dworkin. Allowing reason-based discretion maximises the type of evaluative criteria that can possibly be brought into play in market choice.

This, while it begins to give us an insight into what reason-based discretion is in a market context, does not, however, tell us how reason-based discretion might be legitimately restricted. It is unlikely to be compatible with Dworkin's thesis that reasons associated with persons' race or religion, for example, are always a legitimate basis for reasons-for-choice in the marketplace. So, while it is beneficial for reason-based discretion to be unlimited, it intensifies social competition for scarce resources, it should also be recognised that some limitation is necessary if the market system is to avoid being sullied by unwanted biases and crooked tastes. This does not help, of course, to determine where the line should be drawn, between acceptable and unacceptable preferences, but it does help to draw the relevant distinction between limited and unlimited reason-based discretion. The distinction is an important one for what I want to say about the reflective capacities of the market participant, and the next section attempts to specify it further.

§2.3 Reason-Based Discretion in Markets and Hierarchies

A recurrent debate in the field of what is often called "transaction cost economics"⁵⁴ concerns what distinguishes, in a precise way, market-based methods of organising production from production conducted via hierarchical organisations. An obvious distinguishing feature is the extent to which immediate managerial control is maintained within a hierarchy, in which managers have the power to directly command subordinates, and the extent to which managerial control is diluted in a market-based relationship.

Alchian and Demsetz (1996) claim to have identified a weakness in the view, commonly accepted among transaction cost theorists, that the long-term labour contract is employed by firms because of the managerial command it affords. The ability of a worker to leave the firm or, alternatively, be sacked for non-compliance with demands made, makes the employer-employee relationship more of a market-type relationship than the common view allows. The degree of command is equivalent, say Alchian and Demsetz, to that associated with arm's length, or market-based, transactions. It is not important to become embroiled in the details of that debate, but there is something pertinent to my argument to be

⁵⁴ Coase (1937) is commonly recognised as the groundbreaking paper in this area. The transaction cost approach to the study of industrial organisation was developed later, prominently, by Oliver Williamson (1975, 1998, 2000, 2002), and is often placed under the banner of "new institutional economics". Good summaries of Williamson's influential approach are Williamson (1994, 1996). Essentially, in his 1937 paper Coase set himself the task of explaining the existence of firms. He concluded that the hierarchical control they afford removes the costs of market transacting that would be incurred were all employer-employee relations to be conducted on a market, or "arms-length", basis. Williamson extends the theory in a number of directions including a consideration of the impact of the degree of "asset specificity" on the firm-supplier relationship.

had from considering a possible location of the divide between market and hierarchical methods of organising production.

Consider the expert employee, an employee in a firm who has the recognised expertise to make independent decisions concerning, let us say, choice of supplier, for an important material input. Responsibility for the choice, let us imagine, has been devolved to this individual because of his expertise. It is natural to suppose that there is an implicit understanding that the reason the expert has for selecting suppliers and the reason the firm has for asking him to make the selection are compatible. The reasons of the expert employee and the employer need not be identical, however. What is often placed under the rubric of the “principal-agent problem” is the insight that employer-employee motivations might not coincide. What is important for the employer is that whatever reason the employee has for making the choice should further his (the employer’s) aims. That is, the employee’s preference ordering should be one that *could have been* informed by the employer’s reasons for making the choice among potential suppliers. This gives the impression that *complete* reason-based discretion can be granted to the employee, but this is not the case. It is, of course, possible that the employee might choose for a reason that is not compatible with that of the employer, its incompatibility lying in the fact that the preference ordering is not one that could have flowed from the employer’s own reason for making the choice. The employee might be inspired by arguments in support of fair trade, for example, to opt for a higher cost supplier, arguments that do not, let us assume, similarly inspire his employer. The drawback for the employer, then, of complete reason-based discretion is that reason compatibility might be lost. Moreover, if

reason-based discretion is complete there will be no mechanism by which the employer can get the expert employee back on track. The loss of hierarchical control, which complete reason-based discretion implies, will result in a loss of control over the preference ordering of the employee. Recalling the description of the components of preference formation (section 2.1, above), we can say that the expert employee can be permitted complete evaluative discretion, complete discretion over what evaluative criteria to apply – therein lies the importance to his employer of his expertise – but, in order to preempt the possibility of loss of control over final preference, reason-based discretion *cannot* be unlimited. That is, the expert employee can be granted evaluative discretion, indeed can be granted complete evaluative discretion, but he can only be granted this given a prior limitation on the reasons he can apply in making the relevant type of choice. Hierarchical organisation is incompatible with unlimited reason-based autonomy, on the part of the employee. Full autonomy can extend to choice of evaluative criteria but only on condition that the reasons a person applies fall within a permissible range.

In contrast, a market system is compatible with *unlimited* reason-based discretion. The compatibility of the reasons a person has for choosing as she does and the reasons of a higher authority need not be an issue.⁵⁵ While the “because”

⁵⁵ I should say that I don’t think I have answered Alchian and Demsetz (1996) in what I say here. The ability to sack an employee (or adjust her contract (Williamson 2002, 185)) is one way to exert control on what she does. The effort at control is, in the terms I have used in the text, an effort to control the reasons she regularly applies. Were the relationship a contractual market one then other mechanisms would be employed, including possible termination of the contract, to exert control. What is different in the two types of relationship (market and non-market) is the way control is exercised. In this

of personal choice (Cf. §2.1) is necessarily limited in a hierarchy, it need not be in a market, quite simply because what characterises personal choice in the latter is an absence of formal managerial superintendence (of course, there might actually be superintendence in unequal relationships, for example). Hence, what in Dworkin is viewed as a valuable property of a market is, further, a distinguishing feature of a market, in that it is compatible with market organisation but not with organisation that has a hierarchical structure.

However, it is not necessary for market participants to have unlimited reason-based discretion in order for the system of exchange they are involved in to qualify as a market. While Dworkin would argue, I think, were he to present things in these terms, for reason-based discretion, he would acknowledge the need for limitations on such discretion in order to root out, for example, crooked tastes. And he would not consider that the existence of such limitations is incompatible with market-based exchange - limited discretion over personal reasons-for-choice is compatible with market participation. For example, legislation designed to combat discrimination in the selection of job candidates aims, in part, to outlaw types of reason for choice. A person who takes this on board in her recruitment of personnel is not thereby disqualified as a market participant. Indeed, while I have argued that Cohen's preferred restriction on market conduct does not follow from Rawls' principles of justice, a person may well apply such a restriction on his or her reasons as a matter of personal, moral choice. This self-imposed limitation of reason autonomy would not interfere with her status as a market participant. This

way we can see that Alchian and Demsetz have in their sights two types of relationship – one hierarchical and one market-based. What I consider in the text is a contrast between two types of organisation per se – one hierarchical and one not.

may seem to make the notion of unlimited reason-based autonomy redundant, or relevant only for a hypothetical market system in which only basic market rules of exchange are in force, those preserving property rights and preventing fraud. But, as we will see, the distinction between unlimited and limited reason-based autonomy, far from entailing the redundancy of the former notion, is a helpful one.

To see this, note that we can distinguish between a market understood as a system of voluntary exchange, shorn of any but minimal institutional restrictions, designed to make stable exchange possible, and a market understood as subject also to a range of market-oriented policies – designed to regulate competition, proscribe harmful goods, and so on. This sort of distinction is in, for example, Lindblom (2002):

Although not all societies embrace or contain a market system, all existing societies make use of markets...Whenever people frequently pay other people to do something – sing a song or dig coal – those interchanges constitute markets. Yet despite the commonality of such interchanges in Maoist China or the Soviet Union, these societies were not called market systems, because a market system exists only when markets proliferate and link with each other in a particular way...specifically to organize or coordinate many of the activities of a society. (Lindblom, 2004, 4).

Similarly, we find such a distinction in Max Weber's *Economy and Society*:

A market may be said to exist wherever there is competition...for opportunities of exchange among a plurality of potential parties. Their physical assemblage in one place,

as in the market square,...only constitutes the most consistent kind of market formation. (cited in Swedberg, 1994, 265).

This, in addition to Weber's later contrast (noted in Swedberg, 1994, 266) between a market and the regulatory framework that conditions personal action brings out the type of contrast I have in mind.

Let us call a market understood as being constituted by voluntary exchanges, along with the enforced restrictions that facilitate such exchanges, a *market per se*. Were a person to be involved (hypothetically) in such a basic market set-up, she would have unlimited reason-based discretion. There are, in such a set-up, no institutional restrictions on the reasons people might bring to bear in their personal choices. So we can distinguish between the market a person is actually engaged in and the market *per se*, the market set-up in which she might, counterfactually, be engaged, a market in which limitations on reason-based discretion are (counterfactually) silenced.⁵⁶ A market participant can, then, distinguish between the market she is involved in, replete with higher-order⁵⁷ constraints, and the market *per se*, the market shorn of its higher-order,

⁵⁶ It might help to note that two market systems can be identical, understood as markets *per se* - in both there may be rule-governed, voluntary exchange – while being very different in their less basic institutional characteristics. In one market system there might be a relaxed, “Austrian” approach to monopoly power, while in the other institutional rules condemn any form of market domination.

⁵⁷ It is useful to introduce, at this point, the term “higher-order” to describe the restrictions and permissions associated with market-oriented policies. What is “lower-order” are the more basic provisions, that make trade possible. I will also draw on the contrast between the contingency of higher-order rules and the necessity of lower-order provisions.

institutional properties. An immediate corollary of this is that higher-order market rules are contingent and, being so, are *characteristically open to critical review* - being contingent their practical efficacy is conditional on their being warranted.

Now, recall that in Rawls' theory, as he presents it (Cf. §1) everyday market activity is informed by comprehensive doctrine, while the institutional backdrop, that determines, inter alia, the nature of competition, redistributive mechanisms, and, which goods are deemed tradeable, is informed by neutral, behind-the-veil reasoning. On one side, everyday market activity is informed by comprehensive doctrine. Attitudes to consumerism, reluctance to eat dead animals and to purchase sweatshop goods, and so on are influenced by religious and group affiliations, and what persons see as personal crusades. On the other side are commitments to policies, and their associated implications for personal conduct, that are supported by a mode of reasoning that does not engage the concerns that emanate from comprehensive doctrine. To make the next step in my argument I want to notice that this seemingly happy schism is hard to maintain in the face of policies that have implications for practical reasoning in the personal domain. A policy intended to root out unwarranted discrimination in recruitment, for example, could be interpreted as requiring, on the part of the market participant, the adoption of a rule designed to disallow, prior to actual deliberation, certain reasons for choice. However, when it actually comes to selecting among job candidates the application of the rule will not alone serve to moderate the criteria the participant applies, *in the act of* selection. A person, for example, whose choice is influenced by the ageist norms that prevail in her society will have to tussle with the criteria she would readily apply were there no policy intended to

root out ageism. Similarly, a policy designed to break prejudice in consumption patterns will have to work in part on attitudes if the policy is to have practical effect at ground level. Hence, it is implausible to presume a happy separation (contra Rawls) of commitments between those in support of basic institutions and those that directly influence personal choice. So, institutional rules will, in many (important) instances, influence the application of evaluative criteria in personal, market-based choice. (Relevant discussion is in Nancy Fraser (1999). See addendum I).

Now, recall that everyday, market-oriented choice is informed by comprehensive doctrine (in Rawls), and add to that the fact, just observed, that higher-order market rules have (in many cases) implications for actual market-based personal deliberation. This entails that a person's deliberation as to whether they should accept, as a factor in their personal choice, the evaluative criteria associated with an institutional rule, must be informed by their comprehensive doctrine. This is because the practical implications of the institutional rule are considered at ground level, from the perspective associated with the everyday stance of the market participant, and that stance is informed (in Rawls' well-ordered society) not by neutral reasoning but by comprehensive doctrine (the central point of §1). In order to bring out what is potentially achieved here, I note a seeming difficulty. If the justness of the implications for personal choice of an institutional rule is assessed from within comprehensive doctrine then its justification will, according to the objection, be assessed from an inappropriate point of view. While the critique proceeds in a way that engages personal outlooks and the concerns that emanate from them, this is bought at the cost of removing a

point of view that transcends comprehensive views. Without such a transcendent viewpoint, social agreement as to the content of public rules, and their underlying principles, will only be achievable if all can agree from within their comprehensive doctrines, and this will lead us, surely, into the type of “comprehensive liberalism” Rawls is concerned to avoid.⁵⁸ I have to argue, then, if the conception I am defending is to be friendly to Rawls, that presenting the possibility of dissent in this way does not necessitate partial, substantive agreement among otherwise competing conceptions of the good. I can, I think, argue this, and the form the argument takes delivers the picture of critique that, I think, secures a better (more realistic) representation of the problem of social stability.

⁵⁸ He says (Rawls, 1999, 360): “one deep divide between conceptions of justice is whether they allow for a plurality of different and opposing, and even incommensurable, conceptions of the good which is to be recognized by all persons, so far as they are rational...Plato and Aristotle, and the Christian tradition as represented by Aquinas and Augustine, fall on the side of the one (rational) good...The presupposition of liberalism...as represented by Locke, Kant and J. S. Mill, is that there are many conflicting and incommensurable conceptions of the good, each compatible with the full autonomy and rationality of human persons.” Hence, the liberal inclination is to avoid an appeal to a single conception of the good, but it is, says Rawls, the mistake of “comprehensive liberalism” to suppose that, among competing conceptions of the good, there is a common baseline from which to work towards political, if not doctrinal, harmony. He accuses Kant and John Stuart Mill, and his younger self, of comprehensive liberalism.

§3 The Argument Adopting a “Metapreference” Approach

A good way to present the argument is to adopt, for purposes of exposition, what has been called a "metapreference" conception of the process of choice. (See addendum II).⁵⁹ Consider a person's choice set, $C\{x, y\}$, being the list of (here two) objects that the person can form a preference over, and select. They might be candidates for a job, or two potential housebuyers, recalling the earlier examples. P_1 and P_2 are two preferences over x and y that the person may hold, xP_1y and yP_2x . There is, further, an institutional rule whose effective implementation requires that people moderate the evaluative criteria they would otherwise readily apply in market choice. An appropriate moderation of evaluative criteria (from the point of view of policy) is reflected in P_1 , while P_2 is the ranking that reflects the unmoderated application of the personal evaluative criteria of the market participant. We have, then, if a person chooses in accordance with P_1 , $C\{x, y\} = x$, and if P_2 then $C\{x, y\} = y$.

Let E denote the evaluative criteria that support P_1 . According to the metapreference interpretation of choice, that P_1 holds sway can be represented as $(P_1)P(P_2)$, that is, preference ordering one is preferred to preference ordering two. So, if E is accepted by the market participant then $(P_1)P(P_2)$ which entails $C\{x, y\} = x$. Letting E' denote the personal evaluative criteria of the market participant we have if E' then $(P_2)P(P_1)$ and $C\{x, y\} = y$.

If we omit the intermediate preference rankings we have, more simply,

if E then $C\{x, y\} = x$ and if E' then $C\{x, y\} = y$

⁵⁹ The way things are laid out here is influenced, most directly, by Sen (1997).

Now, this metapreference construction stops at the level of evaluative criteria. However, each criterion has an associated justification. A justification for E would be defended by the authorities who would like to see E reflected in personal choice. The justification for E' (personal evaluative criteria) is grounded in the comprehensive doctrine of the agent. So, let J denote the justification for E and J' the justification for E'. Note that each justification can be embedded in a "because.." clause, of the form, "evaluative criteria E (or E') ought to be applied because...". Note also that, in order to represent the reason-based autonomy that, I have argued, ought to be ascribed to the market agent, we can represent a higher level still at which the because clauses, rather than the evaluative criteria they support, can be compared. Indeed, that it facilitates a representation of reason-based autonomy is the chief virtue of using a metapreference approach in this argument. Now, one way to express the implications of the distinctions I have identified in connection with market participation is to say that counterfactual autonomy with respect to "the because of choice" is what is secured for the market participant, by virtue of his participation in a market (non-hierarchical) system. This makes live in the comparison of justifications the *possibility* that it is the justification that reflects personal concern that will inform choice. At the same time we avoid the objection that this is achieved at the cost of presenting the adjudication from the perspective of comprehensive doctrine. That the objection is avoided can be seen from the way in which in the metapreference structure the two justifications are brought into view. They are compared from a higher vantage point that need not be seen as itself informed by comprehensive doctrine. Neither

need it be characterised as neutral. I am, in fact, suggesting that counterfactual autonomy with respect to the because of choice equips the person with a criticality that transcends the neutral/non-neutral divide, such as that presented by Rawls.⁶⁰ Persons are capable of a critical appreciation of competing justifications, and their grounds, that does not presuppose a particular normative stance, a criticality that does not require us to see persons' thinking as falling on one or the other side of a divide between neutral and non-neutral modes of evaluation. Moreover, competing justifications (personal and institutional) are in view and vie with each other in an assessment of the merit of existing policy. This conception keeps in view the personal concern that motivates dissent because it does not require the agent to transcend her everyday outlook in order to assess institutional rules.

§4. Implications for Cohen's Argument?

At the end of chapter three I rejected a response Cohen might make to the accusation that he fails to pay proper respect to the division between neutral and non-neutral standpoints in Rawls. I said that Cohen might argue that the principle restricting a certain type of personal reward is delimitative, in that it rules out, prior to actual choice, a consideration of effort-reward combinations where rewards are pure incentive payments. Being delimitative, Cohen might argue, they do not invade the personal domain whose practical concerns are shaped by comprehensive doctrine. This, I argued, will not work, as what Cohen wanted to show was that a *personal* principle is entailed by a commitment to the difference

⁶⁰ The seeds of this type of view are in Sen whose metapreference model of moral choice has the agent opting for a particular mode of moral thinking from a perspective that lies outside of that mode.

principle, i. e., a principle that informs choice, as it were *on location*, in the marketplace. A delimitative normative principle is not personal in that sense. This way of rejecting an appeal to a delimitative-personal distinction is, perhaps, brought into question by the discussion of this chapter. A delimitative normative principle is a person-level analogue of an institutional rule. That is, the person is viewed as having, so to speak, a personal constitution⁶¹ that rules out anterior to any actual choice taking place a consideration of (staying with the Cohenist case) the pure incentive structure of a possible reward package. The constitutional character of a delimitative principle contrasts well with, what I am calling in this context, a personal principle, because the latter can be pictured as emanating from comprehensive doctrine, while the former is supported by neutral thinking. This preserves well, in a schematic way, the compartmentalisation that appears to follow from Rawls' distinctions, and serves, it seems, to rebut Cohen. However, the discussion of this chapter appears to break down this seemingly neat division between delimitative/neutral principles on the one hand and personal/non-neutral principles on the other. I have argued for a conception of reflective critique that takes in in a single sweep justifications for institutional rules, not tied to comprehensive doctrine, and justifications for personal evaluative criteria that are so tied. My defence of Rawls against Cohen relied on the notion that everyday, market-oriented choice is informed by comprehensive doctrine, and that imposing moral requirements (derived from institution-level principles) in this sphere was an indefensible intrusion. If agents are able to take in, in one mental sweep, justifications for both institution-level principles and justifications associated with

⁶¹ A similar notion is in Vanberg (1994, 21).

comprehensive doctrine, then why can't they appreciate and accept the entailment Cohen defends, without this being seen as an invasion of a privileged, personal space? In answer to this, I suggest that the fact that a person has reason-based autonomy as a market participant prevents a moral principle (derived from reasoning offered in support of institution-level principles) overriding a market participant's capacity to reflect on and apply self-chosen evaluative criteria in the marketplace.

§5. Conclusion

The central claim of this chapter has been that an adequate conception of embedded critique – adequate because it can help secure a realistic representation of a stable society – can be derived from a characterisation of the market participant. According to this characterisation a distinguishing feature of a competitive market is the *reason-based autonomy* afforded its participants. That is, the reasons for which agents choose is a matter of personal discretion. It follows from this characteristic of market participation that it is inappropriate to represent the agent as someone for whom reasons for choice (in the market context) are imposed in a way that overrides that discretion. If that discretion is overridden then what we have in view is not a market but, as I have argued in the chapter, a hierarchical system. It must be the case, then, in order to preserve that discretion, the discretion associated with reason-based autonomy, that a person who acts in accordance with an institutional requirement intended to restrain market conduct (for example a piece of legislation intended to root out discrimination in the marketplace) must be understood to do so in a way that

preserves that discretion. That is only possible if the reason associated with the institutional rule is brought into view as a reason that *might not* be applied, that might not, if the market agent so decides, be the reason that informs his choice. For the compliant dissenter, someone who reluctantly chooses in accordance with the reason associated with an institutional requirement, for the relevant type of discretion to be preserved in her market-oriented choice it must be the case that the reason associated with the (reluctantly accepted) institutional requirement is brought into open conflict with the reason associated with personal concern (the concern that motivates the dissent). This must be the case if it is to be a live possibility that the agent might act from personal concern, and so shun the institutional requirement. For this vying of competing reasons to take place in the practical reasoning of the market participant it must be the case that there is a vantage point from which both types of reason, and their rational support can be brought into view. The vantage point both transcends the points of view associated with neutral and doctrinal perspectives and keeps in play, in deliberation, the motivating concern that fires the dissent. The market participant, then, ought, if the discretion associated with reason-based autonomy is to be preserved, be pictured as capable of a mode of thinking that is both critical, because it brings into view competing reasons or choice, and embedded in that it keeps in play the socio-cultural bases of personal concern.

Addenda

I

Here I offer further support for the view that policies associated with basic institutions ought to be viewed as having, in some important ways, implications for personal choice in the marketplace. I draw on the work of Nancy Fraser who defends what she sees as an alternative paradigm to the type of approach she associates with Rawls.

In her "Social Justice in the Age of Identity Politics" (1999) Nancy Fraser distinguishes between a concern for economic and social justice that focuses on the distribution of resources, or a derivative metric such as welfare, and an approach that emphasises the role of "recognition" imbalances in generating and supporting inequalities. A just society is, on a recognition account, one that recognises, and accommodates differences among personal identities: differences of ethnicity, sex, race, and gender. I take this to be for two reasons. Recognising differences associated with, for example, physical characteristics, can provide reasons for special treatment (for example, adjusting public spaces to enable access to people with disabilities can make their preferences "authentic" (Macleod, 1998, 38)). The explicit recognition of difference can, on the other hand, provide a basis from which to argue for equal treatment, for those differences not to be taken into account inappropriately. Apposite are employee selection procedures (referred to in the text) that are indifferent to class or sex. Talk of claims based on personal identity are, in large part, orthogonal to a discussion of resource or welfare inequality. It is possible to examine the way social roles track gender, for example, quite independently of how the social division of tasks shows up in material holdings or, say, personal expectations.

However, Fraser has something interesting to say about the seeming disparity of the two approaches. First, it is important to her that a recognition approach is not limited to "identity politics", to the "affirmation of difference" (Fraser, 1999, 27) – taking this to be related to the realm of understandings, or "significations", as Fraser puts it - but includes "currents dedicated to righting gender-specific, 'race'-specific, and sex-specific forms of *economic* injustice." (ibid. My emphasis.). Hence, an emphasis on recognition brings to the fore causes of inequality related to, for example, racial or sexual prejudice. And, importantly, recommends cures that call for cultural change rather than material redistribution. As Fraser puts it:

The redistribution paradigm focuses on injustices it defines as socio-economic and presumes to be rooted in the political economy...The recognition paradigm, in contrast, targets injustices it presumes to be rooted in social patterns of representation, interpretation and communication. Examples include cultural domination, non-recognition, and disrespect...In the redistribution paradigm, the remedy for injustice is political-economic restructuring. This might involve redistributing income...In the paradigm of recognition, in contrast, the remedy for injustice is cultural or symbolic change. This could involve upwardly revaluing disrespected identities." (Fraser, 1999, 27).

Fraser's comments are of particular interest here because of what they imply for the relationship between policy and personal choice. It must be the case that recognitional concerns are intended to feed into personal market conduct in Fraser's thesis. Were exchange to be infected with the sort of recognitional imbalances she mentions then it would be appropriate, indeed, surely necessary to have that conduct change – through attitudinal change, or through compliance. And for this, it is necessary, if the change is to be policy-based, to presume policy to effect a shift in personal modes of reasoning.

Supporting this view, Fraser (p.31) identifies a type of inequality whose origin is both economic and cultural, circumstances in which "groups...suffer both maldistribution and misrecognition *in forms where neither of these injustices is an indirect effect of the other, but where both are primary and co-original*"; noting that in order to address the inequality both redistributive and recognitional remedies will be required. For example, giving resources to a minority in order to increase their average income level, even if the resources are targeted for the acquisition of marketable skills, will not improve their lot if they are routinely turned down for jobs on the basis of their identity (Cf. Fraser, p.32). Similarly, access to schools with sound academic reputations is commonly based on an ability to buy, consequently expensive, local houses, but such clusters often mirror concentrations categorised according to ethnicity. Hence, on Fraser's account, sensitivity to recognitional issues is not a side concern, but ought, rather, to permeate the market participant's thinking in a market context. Policy ought, on her view, to influence attitudes, and so practical reasoning at a personal level, and involve a commitment to institution-level rules. Helpfully in the present context, Fraser would argue that she is not alone in supposing these two types of concern at play in egalitarian thinking. She states that, while their approaches are "unsatisfactory" (p.36), Rawls, Dworkin, and Sen have recognitional concerns:

John Rawls..at times conceives 'primary goods' such as income and jobs as 'social bases of self-respect', while also speaking of self-respect itself as an especially important primary good whose distribution is a matter of justice. Ronald Dworkin, likewise, defends the idea of 'equality of resources' as the distributive expression of the 'equal moral worth of persons'. Amartya Sen, finally, considers both a 'sense of self' and the capacity 'to appear in public without shame' as relevant to the 'capability to function'.. (1999, 49, n.7, referenced at p.36).

I appeal to common sense in the text for the view that institution-level policy will, at least on occasion, influence personal choice. The type of theory defended by Fraser offers more theoretical support.

II

A metapreference account of moral, more generally reflective critical, reasoning with a view to personal choice, attributes to the agent a preference hierarchy. On this model the agent is presumed to have a range of preferences, some of which may present competing rankings of choice objects. Practical decision then involves a “second-order” decision regarding which of the “first-order” preferences to have determine choice. The metapreference conception appears in a number of largely unconnected literatures, in different forms. Seminal papers are those of Sen (1974) and Frankfurt (1971) in which lower-order preferences are quite explicitly up for examination; but an earlier, more implicit, version of the account can be found in John Stuart Mill (1972/1863). Richard Hare notes his debt to the way in which the vagaries of character and its rational control feature in Aristotle and Plato, and his own moral theory (esp. 1981) itself yields to a metapreference reading. Hirschman (1985) considers the application of hierarchy in choice to economics in a number of ways, including worker motivation. The central idea is to introduce an element of reasoned control into the economist's picture of individual choice, so moving away from the 'black box' passivity ordinarily associated with agents' behaviour. In the work of Elster (1985a) and Schelling (1984) (see also Thaler and Sheffrin (1981)) akrasia is given a treatment that uses higher order preferences to differentiate between what the individual would do were she to successfully restrain herself and how she orders things according to her more pressing desires. George (1993) considers the implications of the model for economic policy, and (George (1998)) contrasts metapreference and multiple-self approaches to modelling irrationality. A review of some of the specifically economic aspects of the metapreference approach is Drakopoulos (1994).

It will be helpful to be clear about what I understand by a metapreference model of reflective evaluation, or, at least, what interpretation of it is in play in the text. And for this I begin by responding to comments that Timothy Brennan makes (Brennan, 1989, 1993) about the character of the metapreference approach, and its theoretical worth. Brennan, whose focus is the conception's would-be role in economics, sees it as a wasteful extension to the orthodox, single-layered notion of preference. Brennan refers to Amartya Sen's (1982a, 1982c – see also 1982b) "model of morality" (Sen, 1982b: 75) as one according to which

higher-order preferences...ought to be followed in situations where actions based directly on satisfying interests would make everyone worse off. (Brennan, 1989, p...).

Brennan says also of Harry Frankfurt's classic *Freedom of the Will and the Concept of a Person* that, according to that paper, hierarchies of preference "reflect different levels of moral importance" (Brennan, 1989: 192); and that,

Using Frankfurt's framework, stealing would be unethical because the preference not to take something belonging to another is a higher-order preference arrived at through moral reflection." (ibid.).

This is, I think, a mistaken interpretation of those accounts.⁶² A consideration of where Brennan goes wrong helps to bring out what I take to be the more instructive account of a metapreference conception I draw on in the text.

What Brennan is associating with Sen and Frankfurt, it seems, is a preference structure in which preferences are categorised as higher or lower order according to their type. In the case of Frankfurt, for example, moral preferences are higher-order because they are subject to reflective endorsement. This way of seeing what is involved in metapreference is suitable to aspects of John Stuart Mill's moral philosophy, and Brennan applies it to Mill. But even in this case Brennan's reading is, arguably, less instructive than the alternative I take to be the correct rendition. It is helpful to look at what he says about Mill. On Brennan's interpretation, the aspect of Mill's moral philosophy that can be given a metapreference reading is his distinction of pleasures, and associated preferences, as more or less sophisticated:

⁶² Mark Lutz (1993) also thinks Brennan's account mistaken.

Mill argued that the pleasures must be judged not just on “quantity” but on “quality”. The pleasures that engage the “higher faculties”, such as intelligence or education, are the more valuable. (Brennan, 1989: 191).

As is well known, contra the views of his father, James Mill, and the position defended by Jeremy Bentham, John Stuart Mill saw importance in a distinction of merit between the higher pursuits of, say, opera, and more base pursuits, such as, for example, watching mud wrestling. Acknowledging the distinction in earlier writers sympathetic to his position Mill points out,

however, that utilitarian writers in general have placed the superiority of mental over bodily pleasures chiefly in the greater permanency, safety, uncostliness, etc., of the former.....rather than their intrinsic nature(..)..It is quite compatible with the principle of utility to recognise the fact that some kinds of pleasure are more desirable and more valuable than others. (Mill, 1863: 8).

And the corresponding distinction in terms of preferences is between those of, say, the savant and those of the couch potato. The temptation here, in reading a metapreference model into Mill’s work, is that we focus exclusively (as Brennan does) on the quality of the preferences themselves, viewing the preference-hierarchy as importantly one of kind, when what is correctly viewed as higher-order here is a preference understood as a reflective ranking of possible preferences; a ranking that brings into view the evaluative criteria supporting the preferences, or (as in Mill) what they mean for one’s character, and so on. When Mill goes on to describe the competent judge, one who is qualified to offer a “decided preference” (ibid: 8) between two pleasures, he has him bring possible pleasures, (and associated preferences, I should add) into view. Thus, preferences of supposed greater sophistication and those more base are equally lower-order. What is higher-order is a considered assessment of their merits.

As I understand it, then, a metapreference account of evaluative deliberation, as witnessed in Mill, sees choice as taking into account possible preferences, these latter being equally lower-order. And the higher-order preference reflects a comprehensive evaluation that takes into account the reasoning in support of the lower-order preferences, such as their perceived personal or social consequences. That this is the proper place of metapreference in Mill’s moral theory is supported by McPherson (1982). According to the latter one way in which Mill argues for the superiority of the higher pleasures is that,

those who are genuinely free to choose their plan of life...display a strong tendency to choose a life of higher pleasures...In the language of modern preference theory, such propensity would be characterised as a “metapreference”: a preference concerning the type of preference structure one wants to have. (McPherson, 1982: 11).

The view also finds support in Martin Hollis’ emphasis (Hollis, 1994) that the political value of individual liberty, for Mill, resides in the room it makes for self-direction, both in given choice-scenarios and in willing adjustments to character and capacities over time. This is a well documented aspect of Mill’s work and will include, I think we can take it, instances in which what are up for adjustment are the untutored preferences a person actually has. Giving Hollis’ emphasis on self-direction a metapreference interpretation we would distinguish possible preference-types as lower-order and the decided preference for, say, those associated with a more sophisticated personality, as higher-order.⁶³

What is “higher-order”, then, in a hierarchical conception of evaluative-practical thinking is a decision concerning which possible lower-order preference ranking to have inform actual choice, and not, as Brennan would have it, preference types of greater

⁶³ Making a related though different point, Hare (1981:25), whose own moral theory displays a metapreference structure, refers to Mill as exemplar of the type of approach that sees the source of our moral being as grounded in a capacity to exert intellectual control over the type of emotional response we share with other animals. Mill himself says (1863:54), "It is common enough certainly, though the reverse of commendable, to feel resentment merely because we have suffered pain; but a person whose resentment is really a moral feeling, that is, considers whether an act is blamable before he allows himself to resent it - such a person, though he may not say expressly to himself that he is standing up for the interest of society, certainly does feel that he is asserting a rule which is for the benefit of others as well as for his own." We can see in this quote that the moral individual is one able to recognise two possible preference orderings: one associated with unmitigated emotional response and one with rational control. The two orderings need not conflict in terms of the action they recommend but it is possible they will. The quote has Mill observing not simply a ranking of preference types but considering the role they play in a reflective consideration of what one ought to do. This, again, tells against Brennan's version of the model.

sophistication. While a metapreference reading of Mill is available that ranks preferences according to qualitative criteria, I take the interpretation outlined here to better capture Mill's intention. It is this form of metapreference model that I intend in the text. There is, however, more to be said on this that is useful in motivating the use of metapreference in the text. What is brought into view in a metapreference account need not be limited to evaluative criteria, but the modes of reasoning that support them as well, as the following points out.

We find in Sen an explicit intention to present a conception of metapreference that models moral deliberation. Its structure is, perhaps, most clearly expressed in the following:

Let X be the set of all possible outcomes and π be the set of all possible orderings of elements of X . A moral view can be defined as a quasi-ordering Q of the elements of π ...What does a moral quasi-ordering Q stand for? Various interpretations are possible. It might take the form of a moral desire to have one preference pattern over outcomes rather than another...Or else Q may correspond to one's ranking in terms of praiseworthiness...(Sen, 1982b: 80).⁶⁴

This set-theoretic formalisation, whilst rudimentary, seems, on the face of it straightforward enough as a possible, if stylistic, description of moral choice. Frankfurt, unlike Sen, talks not in terms of orders of preferences but of desires. Peculiar to the

⁶⁴ A quasi-ordering is, following Sen's usage here, a ranking that is reflexive and transitive but may not be complete. One way to contrast rankings is in Sen (1970) chapter 1* but definitions vary across authors. I note also that, in later work, Sen (1995, 1997) has discussed the importance of context in personal choice. While economists make much of the internal consistency of choice, consistency among choices made, and, implicitly, the preferences underlying them, it is commonly the case that the context of choice will influence the choice made. The presence of a child might lead a chocolate-loving adult to choose the apple over the last piece of cake, a choice that may well be reversed were the child not present. There is a strong connection between this type of observation and Sen's earlier work on metapreference. The underlying preference (for cake over fruit) has not changed. What is responsive to context is the adult's higher-order ranking.

human will, he says, is an ability "to form....'second-order desires' or 'desires of the second order'." (Frankfurt, 1971: 6). However, he goes on to say,

Besides wanting and choosing and being moved *to do* this or that, men may also want to have (or not to have) certain desires and motives. They are capable of wanting to be different in their preferences and purposes from what they are..(...)..No animal other than man..appears to have the capacity for reflective self-evaluation that is manifested in the formation of second-order desires. (ibid: 7).

Hence, there is a distinction between second-order desires, as such, and the process by which they are established, what is primitive in this process being a "wanting to be different" (ibid.). A second-order desire concerns first-order preferences, and reflects what a person's higher-order wants are concerning not the objects of the preferences themselves but the relationship between the preferences and her character, her longer-term goals, her religious beliefs, commitments and so on.

In a footnote, Frankfurt notes the following:

In speaking of the evaluation of his own desires and motives as being characteristic of a person, I do not mean to suggest that a person's second-order volitions necessarily manifest a *moral* stance on his part toward his first-order desires. It may not be from the point of view of morality that the person evaluates his first-order desires....Second-order volitions express evaluations only in the sense that they are preferences. There is no essential restriction on the kind of basis, if any, upon which they are formed.' (Frankfurt, 1971: 13).

The pursuit of the rules-of-thumb, say, embodied in a society's moral code may even conflict for an individual with the pursuit of personal gain without us as yet invoking Frankfurt's second-order considerations. What takes us onto the higher level is the individual's reflective evaluation of the desires he has. It is consistent with Frankfurt's model to be in moral turmoil without thereby being understood to be bringing into question the desires one in fact has. What is distinctively human is the capacity to bring moral thinking itself into view, in order, in this case, to evaluate one's commitments. In the text a metapreference is used to achieve just that. A representation of normative

thinking that involves a higher-order comparison of possible evaluative criteria, a comparison, indeed, that does not invoke, or privilege a particular mode of evaluation.

Chapter 4

McDowell, Rawls, and the Market

Introduction

The source of the deficiency in Rawls' representation of a stable society is, I have argued (Ch. 1, §4), following Mark Granovetter, the "undersocialised" character of his conception of reflective critique. A conception of reflective criticism that has the person dissociate herself from her everyday modes of evaluation has her adopt a perspective that extinguishes the concerns that motivate her dissent. Where she accepts that her point of view is faulty, or a minority view, and so a view that ought not to inform her conduct, she has no practical standpoint from within her everyday ways of thinking from which to engage with the motive force of her personal concerns. All that is open to her is a root and branch removal of her motivating concern, and this, again recalling Granovetter, leads to her "oversocialisation". Rather than accommodating the tension that exists in an ongoing democratic society between commitment to existing institutions and a desire to see them changed, Rawls' oversocialised conception, pictures agents as unquestioningly obedient to prevailing policies, practices, and regulations. What is required, in order to remove this deficiency, is a tenable conception of reflective critique that has the agent capable of taking into account both neutral and non-neutral considerations simultaneously. This would enable the comparison of neutral and non-neutral considerations to proceed without the detachment that, I

have argued, sets aside the motive force of personal concerns. Such a conception is, as I tried to show in chapter three, can be derived from a particular characterisation of market participation, a characterisation that emphasises market agents' reasons-based autonomy. On this view, in order to accommodate the role of both comprehensive doctrine and institutional requirement in everyday market choice we must conceptualise market agents as capable of higher-order thinking that embraces both neutral and non-neutral reasons for personal, market-based choice. Being a critical perspective that embraces both neutral and non-neutral points of view, this conception keeps alive the motive force of personal concerns even while they are being subject to critique that engages neutral considerations.

This approach has the important advantage, I think, of supporting a conception of embedded critique that draws on resources from within Rawls' own theory. No metaphysical contentions are imported from outside of Rawls' account, avoiding difficult-to-settle disputes about the character of reflective thinking, and respecting Rawls' belief that political philosophy ought not to rely on metaphysical presuppositions. However, the argument would be only partially successful were it to end here. There are two reasons for this. First, while market choice is an important aspect of everyday life in a market economy it is not, of course, the only, nor the most significant, context of personal choice. And so, in order to claim to have a conception of critical reflection that applies to the Rawlsian citizen as such, something must be said about the implications of my view for the conception of reflective capacities in family, professional, and other social settings. Second, the argument of chapter three draws on an abstract characterisation of agents, qua market participants. Because it relies on abstract

considerations the argument suggests little about how flesh and blood agents in the real world are to come to be equipped with such reflective capacities, and the implications of the conception I have offered for how we conceive their moral thinking more widely.

It is important to extend my thesis in these two ways because the conception of critical reflection I have argued for must be married to a conception of the well-ordered society as an ongoing, organic enterprise in which the commitment to basic institutions people have (and thereby social stability) is engendered, in part, in persons' moral upbringing. This reflects Rawls' approach in *A Theory of Justice* in which (in Part III) he presents a picture of developmental moral psychology intended to complement the abstract discussion of Part I. In part III Rawls' intention is to show how a society whose regulative principles are justified as in Part I can be a stable society - the argument being that a commitment to original position thinking is consistent with a theoretically sound conception of moral development that recognises a universal interest in a stable society and sees moral growth as constituted by gradual enhancements of the capacity for rational reflection. This combination, for reasons I have suggested in chapter one, will not work. The detachment goes too far and, in neutralising the pull of the motive power of personal concern, fails to adequately represent a stable society. What I attempt to do in this chapter, then, is to defend a general conception of moral thinking that is (1) compatible with the conception I argued for in chapter three, and (2) is adequately developmental to be compatible with a conception of the well-ordered society as an ongoing, organic enterprise.

In order to be successful, the general conception of moral thinking I defend must satisfy certain desiderata. It must allow critical reflection on doctrinal commitments that does not neutralise the fundamental convictions, attachments, and beliefs that inform and support those commitments. Neutral reflection should not be characterised in such a way that detachment requires a person to stand outside of his or her comprehensive outlook. At the same time neutrality should be attainable in such a way that agents can be characterised as, in Rawls' sense, reasonable. That is, an evaluative stance that does not draw on the resources of comprehensive doctrine must be possible. I argue in this chapter that a strong candidate for a conception of moral thinking that satisfies these desiderata is that found in John McDowell's moral philosophy.

The chapter proceeds as follows. In section one I begin to prepare the way for McDowell's thesis by describing the tension between social embeddedness and the capacity for social critique in a way that views Rawlsian and Communitarian conceptions of reflective criticism as polar and equally untenable. This way of setting things up views a conception of reflection that accommodates both embeddedness and the possibility of critique as a desirable via media between Rawls and Communitarian positions. In section two, to further prepare the way, I consider the way in which embedded critique might be thought to be achieved in the work of Bernard Williams and Charles Taylor. What are, arguably, weaknesses in their cognate approaches serve to emphasise the strengths, and the salient aspects of McDowell's approach. I consider John McDowell's thesis directly in section three. McDowell, I suggest, offers a way to characterise reflective critique that combines the ongoing vitality of personal

concerns with a capacity to criticise that does not presuppose the regulative role of comprehensive doctrine. In order to understand McDowell's ethics it is necessary to see it in relation to his epistemology. A sub-section brings this out. In section four I draw on Rawls' own words to support a reading of him that is amenable to this McDowellian view. While recognising that Rawls' methodology eschews metaphysics I argue that there is evidence in Rawls for a conception of moral thinking that has, at least, the structure of McDowell's conception. This section is intended to buttress the view that a McDowellian conception of moral thinking can be drawn on to complement the picture derived, in chapter three, from market participation. In section five I return to the issue of market participation and how it might relate to the McDowellian view and its analogue in Rawls' own work, drawing on the contrast between basic and non-basic institutions. Here I argue that the conception of reflective critique I have offered for the market participant can be extended to the Rawlsian agent as such. The argument here recalls the metapreference framework I drew on in chapter three. Section six relates the argument of this chapter to the problem of stability, and section seven concludes.

§1. Embedded Critique as a *Via Media* Between Rawls and the Communitarians

In this section I argue that Rawls' conception of reflective critique (as the Communitarian views it) and the Communitarian alternative, on a defensible reading of prominent Communitarian positions, can be seen as polar and equally untenable alternatives. This serves to underline the desirability of a conception of

reflective critique that accommodates social embeddedness without submitting agents to a type of social determinism that precludes radical critique. I begin with comments on Communitarianism offered by Mark Bevir and David O'Brien in a recent paper.

Bevir and O'Brien (2003) see the Communitarianism associated with Sandel as part of a "social humanist" tradition they trace back to the British idealists, such as Edward Caird and T. H. Green (Bevir and O'Brien (2003, 306)). Their main concern is to argue, indeed, that the Scottish philosopher John Macmurray acted as an important bridge between British idealist and later Communitarian thought. Macmurray, they say (ibid., 312), "defended an analysis of the self as embedded in personal relations." For him, thinking is embedded activity in the sense that it falls under the more general concept of action and ought, as such, to have the personal executive involvement in it emphasised. The unattached, disembodied thinking associated classically with Descartes' *cogito* is thus seen as a misapprehension of the nature of thought and thinking. The essence of one's pure, unmediated being as traceable to the activity of thinking, the basis of Cartesian dualism, is replaced, in Macmurray, with the notion of thinking as something a *person does*, and this makes "the person" a more primitive concept than mere thinking as such. But what, it is helpful to ask, is the relationship between the priority of the person and his or her social embeddedness? D. D. Raphael in his review of Macmurray's *The Self as Agent* says, "Just as the thinking subject of Cartesian philosophy is an abstraction from the practical agent, so the agent considered alone is an abstraction from the person, who is in community with other personal agents." (Raphael, 1959, 268). And avoiding this

second type of abstraction is, Raphael contends, what prevents Macmurray from being a Pragmatist:

Pragmatism, of course, allows for the fact that men live in societies, co-operating and competing with each other. But Professor Macmurray would say that the life of community, of love and friendship, is what *makes* men persons. (ibid. My emphasis).

Hence, what does the thinking is not to be purified of social connections and environmental context but is to be understood as constituted by those types of relations. There is no agent as such that is, or comes to be, clothed in social garbs and so made a communal animal. What is primitive is just the person understood as constitutively social. This is instructive for an attempt to describe McDowell's thesis as a middle way between transcendent and embedded evaluative thinking. I consider the following.

Andrew Cohen (1999, 121) observes that a “social constitution thesis” (his term) is shared by Sandel, MacIntyre, and Taylor (p. 122); being, he says a core aspect of communitarian thinking. And Macmurray's priority-of-the-person response to Cartesianism would seem to fit Cohen's bill. Cohen offers three pieces of evidence for a “social constitution” thesis in Sandel, MacIntyre, and Taylor. The following are the quotes Cohen uses (p. 122):

Community describes not just what they *have* as fellow citizens but also what they are, not a relationship they choose...but an attachment they discover, not merely an attribute but a constituent part of their identity. (Sandel, 1982, 150).

Separated from the *Polis*, what could have been a human being becomes instead a wild animal. (MacIntyre, 1988, 98).

One is a self only among other selves. A self can never be described without reference to those who surround it. (Taylor, 1989, 35).

Cohen's judgement is that a common idea here is that "an individual being on her own cannot be a person, a *human* being, a self." (ibid.). But an individual's being "on her own" seems not to take us to the heart of what is being said in these quotes. First, let us take an individual's being "on her own" as implying physical distance, in maturity, between the individual and her community. None of these quotes can be taken to suggest that physical proximity, in maturity, to community is a condition for personality. While physical proximity is, arguably, necessary for there to be transmission of tradition, it is not the case that being geographically outside of a community later in life (necessarily) undoes the constructive work of culture. Hausheer (1996, 56) notes that for Herder "Germans only remain truly creative among Germans..Icelanders who go to Denmark perish. Europeans in America lose their creative energy..". But while it might be an implication of the "social constitution" of personality that being alienated from her own tradition damages a person somehow, by reducing her creative talents, for example, such distance does not, thereby, challenge the constitutively social character of personality as such. Indeed, it is precisely because of the constitutive character for personality of social setting that distance can be diminishing.

Let's suppose then that "on her own" means something like "just as such", a meaning that doesn't imply physical dislocation, but the agent considered as an isolated item. The communitarian views expressed in these quotes seem to deny there can be a "just as such". In Sandel community is a *constituent part* of

identity, while in Taylor a description of a self is not possible *without reference to those who surround it*. And this is Macmurray's position too, as Raphael emphasises. It is the constituted person that is primitive in his theory, not the agent as such. This is, Raphael notes, what serves to distinguish his position from a pragmatic account (such as that at a later date expressed by Richard Rorty in *Philosophy and the Mirror of Nature* (1980)⁶⁵) that denies a primitive place to thinking as pure, isolable activity. In his paper, Cohen "begin(s) by noticing that when we say a person is socially constituted we imply that her society (or community) contributes to her constitution." (Cohen, 1999, 123). But the second point, that there cannot be a "just as such", helps us to question the suitability of Cohen's account of social constitution for the communitarian enterprise he wants to talk about. The rejection of the "just as such" indicates that it is not part of the communitarian view that society *contributes* to constitution. That way of understanding things would have two items in place: the unmoulded agent – a primitive unshaped item that is not yet a person – and a later moulded version, being the person proper.⁶⁶ The quotes Cohen cites seem to reflect, then, what

⁶⁵ Notable in Rorty is his analysis of the "ocular" prejudice in epistemology that supports the Cartesian view (Rorty, 1979, *passim*). Rorty recommends a displacement of this view in favour of one that does not presume opposition between mind and world – a view that would be favoured also, I think, by Macmurray (and one that is embraced, though not without important reservation by McDowell (1996, 85-86)).

⁶⁶ That Cohen has in mind a two-item conception of individual development in a social setting is evidenced by what he says about Locke. The latter can, Cohen, says accommodate a Communitarian point of view because the infant who begins life *tabula rasa* must have some input so as to develop into a human adult. There is an obvious dichotomy here of infantile *tabula rasa* and later formed personality.

Bevir and O’Brien (2003) see as the communitarians’ heritage in Macmurray: what we can call an “immanentist” thesis concerning social constitution that sees situatedness and personality as coeval.

Cohen goes on to offer four possible interpretations of the social constitution thesis. These are, translated into matrix form, as follows:

	General	Particular
Genetic	A	B
Sustaining	C	D

On interpretation A (the Genetic General interpretation) the social constitution thesis is a claim that the community contributes to the development of the person, qua person. Interpretation B (Genetic Particular) has it that the claim is similarly developmental concerning the *particular type* of person produced. The view associated with C is that community is necessary to sustain or maintain the person qua person, and D denotes the thesis that society is necessary to sustain the person qua particular type. As we have noted that geographical dislocation has no implication for the socialised character of personhood we can dispense, I think, with both C and D. The genetic thesis can also, I think, be dispensed with insofar as it relies on a dichotomy of shapeless and shaped entities that has no place, if I am right, in the communitarian view of things. A rejection of these four possible readings of Communitarianism invokes, Cohen says, an “atomism” in which the

agent floats free of any grounding or ongoing contextual influence. Making his point in terms of “independence” he says:

The first form of independence – call it ‘ontological independence’ – is independence such that none of the social constitution theses is true...This independence is *atomistic* in that it requires the individual be seen as an atom, capable of coming into existence, continuing to exist, and acting, without any social support or input. (Cohen, 1999, 127).

However, the immanentist understanding of social constitution that I am recommending (as a reading of dominant Communitarian positions), that sees the person as inherently social, calls for a rejection of all four of Cohen’s readings that not only does not invoke atomistic independence but is more faithful to Communitarian intentions. These considerations help us to characterise, helpfully, the search for a conception of embedded critique as the search for a middle way between a Communitarian approach that emphasises embeddedness at the expense of a capacity for critique, and a Rawlsian approach that captures critique but envisions it as disengaged.

At the heart of the weakness in the representation of social stability in Rawls is a problematic conception of the person. A defence centred around its artificiality will not serve to insulate Rawls from criticism in this regard because it is not the conception as such that is at issue, it is its implications for, and the role it plays in, his unrealistic account of the bases of social stability. This appears to give the Communitarian position some appeal as an alternative source of a conception of reflective critique. However, given that (if I am right) dominant Communitarians are immanentist, the Communitarian conception merely begs the question how transcendence is possible when persons are fully embedded in their

socio-cultural milieux. Hence, a *via media*, a way to conceptualise critical reflection that represents it as embedded critique will not come from the Communitarian side of the debate. An author whose work is suggestive of a way to effect a middle way between transcendent detachment and socially immersed thought is John McDowell. I discuss his position in section three. First, I consider two authors whose work also suggests a middle way between Communitarian and Rawlsian accounts of reflective critique: Bernard Williams and Charles Taylor. What is arguably weak in their work serves to highlight the strengths of the McDowellian version.

§2. Williams and Taylor

§2.1 Williams

Bernard Williams emphasises what can be called the “dispositional” character of moral thinking. Foundational is his conception of the “subjective motivational set” (Williams (1981b: 102)). This he presents as a more delicate interpretation than that standardly made of David Hume’s theory of practical deliberation (1985/1739-1740, 1998/1751). Contents of the motivational set are not only desires or sympathies understood as purely irrational items, but also,

dispositions of evaluation, patterns of emotional reaction, personal loyalties, and various projects.. (ibid: 105).

We can, Williams thinks, on this sophisticated reading of Hume, understand the desiderative aspects of deliberation as bearing an intrinsic relation to reason,

without thereby challenging their motivational primacy. What is important to note, however, is not only the character of the contents but also the dynamic process by which they are formed, transformed, and extinguished. Indeed, it is with reference to the dynamic character of practical reason that we can see how Williams combines the conceptual nature of motivational factors with their active capacity to motivate. Williams says, “(w)e should not think of (the motivational set)..as statically given.” (Williams, (1981c: 105)). Maturation, learning, reflection will promote the genesis of new wants and attachments and adjustment to existing dispositions. The sophistication required to enjoy, say, conceptual art or modal jazz can be hard won. Williams makes room, also, for alterations to sympathies, etc., that are the involuntary results of experiences, as well as those that are the outcome of reflective thinking. This non-rational aspect of the shaping of dispositions plays a role in Williams’ rejection of models that characterise practical deliberation as explicable, and, indeed, available to self-awareness as an application of principles. As John McDowell approvingly notes (McDowell (1995)), for Williams, deliberation typically brings into play informal heuristics; and this because of the opacity of the operations of dispositional attributes in reasoning what to do. Williams says, “..ethical thought will never entirely appear as what it is, and can never fully manifest the fact that it rests in human dispositions” (Williams (1985: 199-200), and in his reference to the marks of an “individual meaningful life” (ibid: 201) he talks about “its opacities and disorder as well as ..its reasoned intentions” (ibid: 202). Hence, sympathies, attachments, desires are available in reflection, as Williams sees it, but to suppose that they are, as such, what drives practical thinking, as a common rendition of Hume has it, is

crude. These items, as they appear in a piece of practical-evaluative reasoning, possess a practical meaning that makes them inherently sensitive to rational assessment. Moreover, whilst we can articulate what it is that we take to be a reason for us in deliberating towards choice, these articulations do not uncover the factors at play in shaping the practical significance things have for us. What the individual cannot do is adopt a stance that brings into view the factors that possess motive power. Williams, indeed, rejects,

any model of personal practical thought according to which all my projects, purposes and needs should be made, discursively and at once, considerations for me. I must deliberate from what I am. (ibid: 200).⁶⁷

Hence, practical deliberation is taken to be informed, shaped by a complex of inarticulable motivational factors that constitute the dispositional backdrop to practical thinking, and whose dynamic development the individual is in the grip of, but aspects of which he can make available to himself in articulate reflection.

Here, then, is an author for whom critical reflection is ineluctably embedded in the everyday stance of the individual. While reasons can be articulated and endorsed, and desires can be actively suppressed, the dispositional backdrop that constitutes the character of the individual and informs his unreflective and deliberate actions is opaque and cannot itself be self-consciously dismantled. It appears that we have a conception of reflective critique here that offers to keep personal concerns as a live issue for the person while they are

⁶⁷ This echoes Blackburn's point (Blackburn (1998)) that the fatal error of Kantian accounts is making desire a content of deliberation rather than an informant

subject to critical review. Indeed, Rawls (as we will see, in section four) endorses Williams' view. The difficulty is that, while Williams' conception is arguably successful in embedding critique, and so keeping in view the practical salience of personal concerns, unless his view is extended, it is silent on the relationship between the neutrality required for institutional assessment and the partiality associated with everyday commitments. Moreover, we cannot readily adjust Williams' model to accommodate the distinction. What is opaque, and so inscrutable, is the dispositional backdrop that envelopes the individual's thinking, making neutrality, in the sense of a purposefully adopted neutral stance, not an option. Williams provides the means to conceive reflection in such a way that personal concerns retain their vital presence, but he does so in a way that makes it difficult to see how agent neutrality, in the way that Rawls requires it, might be effected. Williams could, of course, argue that neutrality is an option in a society of agents as he describes them. However, what we desire, if we are to stay within a Rawlsian frame, is to accommodate radical difference among comprehensive doctrines, and in a way that does not presuppose substantive agreement among alternative outlooks. Without a capacity to grasp what constitutes the backdrop we are unable to set up discussion among consulting agents in a way that eliminates fundamental commitments.

Hence, Williams offers a promising thesis for present purposes. His conception of reflective critique is potentially very helpful in that the vital significance of existing affections, desires, and concerns is retained in critical review of their practical standing. However, while Rawls applauds Williams' approach, what he fails to realise, it seems, is that Williams conception does not

offer a standpoint, indeed, precludes a standpoint from which existing commitments can be neutrally assessed - where neutrality requires, as it does in Rawls, a point of view that is free of the influence of comprehensive doctrine. Of course, Rawls distinguishes between the artificial conception associated with neutral discussion, and the representational conception he can say he is associating with Williams, and applauding him for. However, if I am correct, what Rawls needs is a conception, of reflective critique, that accommodates both neutrality, in his desired sense, and the living significance of personal concerns. Williams will offer the second half of this, but his conception of the inscrutability of the "motivational set", precludes a reflective detachment that can self-consciously neutralise the practical role of background dispositions, guiding convictions, and so on. Does Charles Taylor's cognate work in this area offer a more suitable model?

§2.2 Taylor

We can begin with Taylor's distinction, in *What is Human Agency* (1985a), between "two broad kinds of evaluation of desire" (ibid: 16). On the one hand we can evaluate a desire in respect of its instrumental efficacy, in that furthering it a person understands herself to be improving her own welfare. A contrast between desires in this respect requires that we take notice only of their relative ability to offer personal satisfaction. The contrast is with evaluation that acknowledges that a wide range of desires typically reflect the characteristics or dispositions of the person whose desires they are. Thus, as with Williams, there is an emphasis on the background conditions that are operative in the genesis of desire. The associated

type of critique of desire will take into account whether the desire reflects envy, say, or malice, or benevolence, the desire's worth being tied to an opinion about the underlying motivations, character traits, and so on, it reflects. This latter form of evaluation, then, permits discussion not only of the strength of a desire, or, say, external constraints concerning the likelihood of its satisfaction, but also its worth. It's the significance of this contrast that I want to try to bring out in what follows.

The relevant distinctions are actually three-fold. Instrumental assessment of desire Taylor refers to as "weak evaluation" (1985a: 16)). This, as has been noted, contrasts desires in terms of, for example, the costs entailed in pursuing them. This evaluation involves a comparison of desires in terms of their objects' "desirability characteristics" (1985a: 17), here their relative costs, and, in the above example, their potential to further personal welfare. Taylor notes that somebody who engages in this type of assessment is

a simple weigher...reflective in a minimal sense, in that he evaluates courses of action, and sometimes is capable of acting out of that evaluation as against under the impress of immediate desire. And this is a necessary feature of what we call a self or a person. (1985a: 23).

It seems, then, that Taylor wants us to distinguish between desire as it features in reflection and what he calls "the impress of immediate desire". This suggests a contrast not only between, on the one hand, weak evaluation and, on the other, evaluation that considers a desire's worthiness (this being what he calls "strong evaluation"), but also a contrast between weak evaluation and a form of assessment that is weaker still – that associated with the "impress of desire". But what is the significance, for us, of this three-fold division?

In his later *Sources of the Self* (1989) Taylor draws similar distinctions. Here we find him distinguishing between "brute reaction" (ibid: 6), a "response that relates to a proper object" (ibid), and, last, a "de facto reaction" (ibid: 7). If I read him correctly, a response to a proper object is an evaluation which involves an articulation of the desire under consideration through a specification of its object. It is a further aspect of Taylor's thesis that desires are accessible only through language and that a division between how one feels and how one articulates that is artificial. In Wittgensteinian vein Taylor breaks down the division between internal response, understood simply as such, and that which is responded to. His classification, "response to a proper object" seems to be intended to convey the ineluctably phenomenal character of emotional response. Indeed, for Taylor, our emotional life is constituted by the verbal account we can give of it. He says,

...it is not as though we started off with a raw material of repulsions and attractions, elans and uneases...On the contrary, human life is never without interpreted feeling; the interpretation is constitutive of the feeling. (1985b: 63. My italics.).

Going back to Taylor's notion of strong evaluation, we can now see that an examination of the worth of desire is on all fours with the conception of the desire itself. That is, we are not to suppose there is desire, as such, reducible to, say, its psychological aspect, and then, on top, as it were, a value assessment of it. Its worth is intrinsic to its character, an examination of its worthiness, being partly constitutive of it. We can see from this that whilst Taylor castigates the utilitarian,

and, indeed, the Kantian, for their exclusive treatment of desires as reducible to their instrumentality, there is, nonetheless, a place for weak evaluation in Taylor's own thesis. It seems consistent with what he says, to note that some desires, in some contexts, are appropriately contrasted in cost-benefit terms, without recourse to notions of worth. However, even this type of desire evaluation is inherently phenomenal in Taylor. Note that he says, in his description of weak evaluation that it involves an assessment of a desire's object in terms of its "desirability characteristics". In this way, weak evaluation can be seen to be a special case of desire articulation that is wholly consistent with his view that articulation is constitutive of desire. The third distinction, as it is made in *Sources of the Self*, concerns "de facto reactions", identifiable, I would argue, with the "impress of immediate desire" in his earlier *What is Human Agency*. Taylor is quite clear that while our emotions are constituted by our articulations of them, the interpretational activity through which desires are known is not to be thought of as floating free of underlying facts, involving the relationship between ourselves and our environment. He says,

...although they (articulations) are constitutive of our feelings, these cannot just be shaped at will by the account we offer of them. On the contrary, an articulation purports to characterise a feeling; it is meant to be faithful to what it is that moves us." (ibid: 64, italics added).

What Taylor wants from his theory is a happy settlement of two requirements on a satisfactory treatment of the role of desire in practical deliberation. He seems, on the one hand, to want to banish non-conceptual activity as an underpinning for what one might say about one's desires, attachments, and so on, because, it seems

to him implausible to "envisage a non-conceptual analogue for..changes in outlook." (ibid: 70). On the other hand he wants to reserve a place for underlying non-conceptual goings on:

What we experience in the dentist's chair is language independent: and we have no trouble imagining an animal having this experience (ibid: 72).

Moreover, he seems – and this is the sense I attach to “de facto reactions” – to want there to be a place in his theory for a conditioning or constraining role for the non-conceptual. That the phenomenal aspect of objects that evoke moral response is indelible in a proper account of moral experience is captured by Taylor's notion of "intrinsic description" (1989, 7). But, as I read him, what provides for Taylor the informational basis for moral experience is a direct, itself unreflective, contact with the exterior world – a de facto reaction. And in this way, the non-conceptual threatens to intrude on what is taken, in another part of Taylor's theory, to be exclusively interpretational. And it seems that we cannot have things both ways. To explore this a little further, it is useful to note an example Taylor gives of a brute reaction – the example of nausea (1989, 7). (A similar example Taylor gives is of physical menace: "We can give its meaning...in purely medical terms" 1985b, 54). Now nausea (in a simple gastric sense) is amenable to a reduction that mentions only natural facts. That the offending bacteria were in the Roquefort need not be relevant to an account of the nausea qua brute fact (though it might be to a prevention of its recurrence). A contrasting "response to a proper object", for example, a father's feelings about his own daughter's forthcoming performance at the school play, is not amenable

to such a naturalist-reductive account. A reduction that mentions, say, only desire as such will leave out much (or all) of what it is about the forthcoming event that makes it a source of special claims on that father's attention; that makes it a source of reasons-to-do specifically for him.

However, this latter type of response, whilst not open to a reduction that leaves out its phenomenal details, is nonetheless, it seems Taylor wants us to believe, grounded in something motivationally true about the father. As quoted already, Taylor says:

an articulation purports to characterise a feeling; it is meant to be faithful to what it is that moves us. (1985b: 64).

Further, while emotions, etc., are available only in language,

this is not to say we can change our emotions arbitrarily by applying different names to them...It is not just applying the name that counts, but coming to 'see-feel' that this is the right description...(O)ur thesis here that language is constitutive of our subject-referring emotions says nothing about the order of causation. (ibid: 69-70).

As I read him, Taylor intends a "de facto response" to play the role of that which underlies what is available only in articulation, understood as an item that, whilst non-conceptual, is not open to reduction. Hence, when he says that his thesis

takes intrinsic description seriously, that is descriptions of the objects of our moral responses whose criteria are independent of our de facto reactions (1989, 7),

the distinction is intended to invoke a contrast between reflective understanding and underlying emotional activity that does not put the former on the side of the conceptual, and the latter on the side of the non-conceptual. In this way he seems to be alive to the potential for an unsatisfactory rift between reflection, on the one hand, and its material, on the other. This is further supported by what he says in his recent comments on John McDowell.

In his recent response to John McDowell's *Mind and World* Taylor draws on the idea of "coping" as it features in the work of Heidegger and Merleau-Ponty. Coping is, he says, the "ordinary way of living and moving around in the world, and dealing with things." (Taylor, 2002: 111). It is, he continues, this "pre-conceptual" engagement with the world that forms the "background" (ibid: 112) of our thinking: "conceptual thinking is embedded in our everyday coping" (ibid);

The mass of coping is an essential support to the episodes of conceptual focus in our lives...(F)undamentally, the background understanding we need to make the sense we do of the pieces of thinking we engage in resides in our ordinary coping. (ibid).

What Taylor wants is a middle ground between the non-conceptual basis of what we take our feelings to be, and the irreducibly conceptual character of our emotions as they receive attention in our reflective moments. This happy midpoint is not strongly suggested by talk of what we share with animals, as in the dentist's chair example, as what impinges on us is here wholly non-conceptual. In his response to McDowell, however, it seems that in the notion of coping Taylor has found what he wants. The underlying interplay of environment and subjectivity, and the dynamic development of the emotional life, something we

saw at play in Williams, need not be taken to be non-conceptual, and therefore not something that intrudes from outside of our articulations of our feelings. What forms the subject matter of our comments on our feelings is not non-conceptual, but pre-conceptual, a notion that allows for it to be understood as material for conceptualisation without itself being either irreducibly conceptual or non-conceptual. What becomes available once we take on board Taylor's suggestions here is an understanding of the *de facto* as a deliverance that is engendered in a dynamic process in which agent and environment are involved in an ongoing subliminal interactive engagement. Taylor says,

When we focus on some feature of our dealing with the world and bring it to speech, it doesn't come across as just like a discovery of some unsuspected fact, like for example, the change in landscape at a turn in the road.....When I finally allow myself to recognise that what has been making me uncomfortable in this conversation is that I'm feeling jealous, I feel that in a sense I wasn't totally ignorant of this before. I knew it without knowing it. It has a kind of intermediate status between known and quite unknown. (2002: 112).

In presenting an argument in which desires are seen as responsive to reason, Bernard Williams, we saw, hopes to offer an integrated conception of moral deliberation that allows for reflection while keeping the ongoing dispositional backdrop seamlessly in the background. Taylor, like Williams, emphasises the need to keep the dispositional backdrop in place. He refers to the "'background picture' lying behind our moral and spiritual intuitions" (Taylor, 1989, 8) and says that "(m)oral argument and exploration go on only within a world shaped by our deepest moral responses." (ibid). These quotes are suggestive of the way in which Taylor, like Williams, sees a morality as an inescapable underlying framework for

deliberation. What he does, however, it seems to me, is to offer a theory of moral thinking that removes from view what Williams takes to be simply opaque. In his earlier work Taylor leaves in the picture *de facto*, underlying motivational factors that are themselves inaccessible to reflection. In later work, however, Taylor (implicitly) addresses this in his adoption of the Continental notion of coping. On this account dispositional alterations, though unreflective, draw in rationality in a way that rubs out reason-independent, motivational factors, as such. As with Williams, however, Taylor fails, I would suggest, to provide a critical standpoint that lies outside of a person's comprehensive doctrine. He presents a satisfying picture of embedded critique that accommodates the vital force of personal concerns while they are under review. He also goes further than Williams towards accommodating the sort of neutrality Rawls needs - in Taylor background conditions are not determining grounds for standing convictions in the way they are in Williams' account. Nonetheless, the background is all embracing in reflection, which fact precludes a reflective stance that has the fundamental bases of doctrinal opinions in view. An account of reflective critique that is more helpful to Rawls is, I think, that found in the work of John McDowell.

§3. McDowell's Moral Philosophy

Recalling section two, the Communitarian (definitively, perhaps, Sandel (1998/1982, 20-21)) seems to think that the only alternative to (what he sees as) Rawls' conception of the autonomous self is a "radically situated subject" who is indistinguishable from his or her passions, interests, values and commitments. But this takes no notice of a third option: a conception of the person as capable of,

what I have been calling, embedded critique, and what can be called, reflecting McDowell's approach, a "Neurathian" position. On this conception a person is understood to be able to bring into view standing convictions, attachments, and modes of evaluation, *while also being* situated within an evaluative framework; one that forms a backdrop to both unreflective thinking, and critical evaluation. To give a flavour of the Neurathian notion as it appears in McDowell, consider the following quote:

Like any thinking, ethical thinking is under a standing obligation to reflect about and criticize the standards by which, at any time, it takes itself to be governed.....Now it is a key point that for such reflective criticism, the appropriate image is Neurath's, in which a sailor overhauls his ship while it is afloat. This does not mean such reflection cannot be radical. One can find oneself called on to jettison parts of one's inherited ways of thinking; and, though this is harder to place in Neurath's image, weaknesses that reflection discloses in inherited ways of thinking can dictate the formation of new concepts and conceptions. But the essential thing is that *one can reflect only from the midst of the way of thinking one is reflecting about*. (McDowell, 1996, 81. My emphasis).⁶⁸

How McDowell defends such a thesis is the subject of this section.

McDowell defends a philosophical position, in his work on moral philosophy, that is, in intention, a moral realism that attempts to overcome the traditional, problematic dualism at the heart of much discussion in moral philosophy; the dualism imposed by the contrast between an evident normativity

⁶⁸ The original context for the image is Neurath (1959, 201): "There is no way of taking conclusively established pure protocol sentences as the starting point of the sciences. No tabula rasa exists. We are like sailors who must rebuild their ship on the open sea, never able to dismantle it in dry-dock and to reconstruct it there out of the best materials."

and an objective, value-neutral sphere. Simon Blackburn's response to Mackie's classic remarks on the location of moral value, for example, fails, in McDowell's opinion because it is cast in terms of a dualism that we really ought not to entertain (Mackie (1990/1977), Blackburn (1985), McDowell (1985)). This theme is expressed in his earlier work in the idea that colour experience, more generally, Lockean secondary “qualities”, ought to be taken to be irreducibly phenomenal. In his later work (McDowell 1994, 1995) the idea that we ought to escape the traditional value/fact dichotomy is expressed via a challenge to the modern conception of nature, and that approach is my focus here. I do not attempt a comprehensive account of McDowell's views, but in order to understand the relevance of his thought to the issues here it is necessary to go some way to explaining his central opinions. (The interpretations and criticisms of McDowell I draw on here include the collected papers in Smith (2002), and McDowell et al (1998)).

§3.1 The Epistemological Basis of McDowell's Moral Philosophy

I begin with a helpful contrast McDowell makes in *Mind and World* between, what he calls, “bald naturalism” (eg. 1996, 57) and “rampant Platonism” (eg. Ibid., 77). As contrasting ways of describing the character of ethical reasoning the former, we can say, is associated with a way of thinking that adopts a stance that takes itself to have modes of normative assessment on display, as objects of enquiry. That is, bald naturalism takes modes of evaluation to be open to analysis from a fully third-personal point of view that does not, as a point of view, invoke the modes of normative assessment themselves as background conditions.

Normativity can, on this sort of view, be apprehended by concepts not imbued themselves with normative content, and that need not be distinguished from concepts that take the natural world, understood as a store of value-free facts, as their domain. Rampant Platonism, on the other hand, conceives normativity as *sui generis*. As such, a mode of normative assessment is only comprehensible from a perspective of active engagement. This has the basic ingredients in a mode of normative assessment fully in play, and so inaccessible to objective, from-outside-of-the-normative-outlook, scrutiny. I should note that McDowell says that rampant Platonism is not to be identified with Platonism, as such (though Plato's rhetorical style, he says, might lead one to think he ought to be), and that Platonism in mathematics, rather than ethics comes closer to being of the rampant type he describes (McDowell, 1996, 77, n.7). Even so, McDowell's imagery here echoes in a helpful way the Communitarian/Rawlsian contrast I adverted to in section one of this chapter.

We will, where bald naturalism and rampant Platonism are all that are on offer, as theoretical approaches in ethics, tend to oscillate between those two polar positions. One is either outside of a way of taking things into normative account (with the bald naturalist), or one is enclosed within it, and so ineluctably caught up in its presuppositions. If the former approach is taken then the background presuppositions that inform the ethical outlook that is under the spotlight are neutralised, indeed, unavailable; and if the latter, Platonist, point of view is adopted then those presuppositions are all embracing. What arises here is the question raised against both Rawlsian and Communitarian conceptions in section one: If the background conditions for a normative outlook are in place while

critique is under way, then it is unclear how an out-of-outlook neutrality can be achieved; if, on the other hand, these presuppositions are neutralised it becomes difficult to envisage how we make any judgements made practically relevant for the fully “thick” point of view, once the neutralised presuppositions are in place. McDowell attempts a *via media*, between, in his work, bald naturalism and rampant Platonism, which effort is helpful in placing in a general light the thesis I have defended that works up a notion of embedded critique from a conception of market participation.

McDowell’s ethics is part of a larger thesis which, at its most general, perhaps, concerns the relationship between reason and nature; a thesis that takes in also his more fully worked out views on epistemology. He makes it quite clear (McDowell, 1994, 89) that the difficulties he perceives to be central in epistemology play a central role too in any attempt to arrive at a satisfactory account of evaluative-practical deliberation (see also, for example, the introduction to the second edition of *Mind and World*, xxiii, §10, and 1994, 205). I present, first, an overview of his thinking which places his ethics within this larger setting. I then move to a consideration of the Neurathian aspect of his view.

At the heart of McDowell’s diagnosis of what he sees as problematic in our understanding of both empirical thought and motivated action, action that corresponds to, perhaps after-the-fact, specifiable reasons, is his contention that modern philosophy (perhaps, better, modernity, in, for example, Giddens’ sense (Giddens, 1990)) is in the grip of an impoverished conception of nature. This conception is, he argues, a paradigmatic component of modern (meaning post-Baconian) scientific method, and effects, he says, a division between mind as the

location of conceptual activity and a natural sphere from which thought is wholly excluded. Moreover, meaning and value are placed on the side of the subject and a “disenchanted” meaningless realm is identified with nature. In an afterword (1996, pp. 181-182) McDowell acknowledges that a scientific self-understanding that denies meaning and value to nature was not unknown to the ancients. However, early Greek atomists, whose “point of view was remarkably like that of modern science” (Russell, 1945, 65) did not, McDowell assures us, see it as compulsory, as in the modern scientific frame of mind, to see *nature* as disenchanted. There was room in their thought for a realm of nature as distinct from the subject matter of science. Indeed, says McDowell, while Aristotle was aware of the atomist perspective he “has no inkling of a perfectly correct thought that we can formulate like this: *if we identify nature as the topic of scientific understanding, we must see it as disenchanted.*” (McDowell, 1996, 182).⁶⁹ Further, associated with this conception of the relationship between thought and world is an implicit commitment, McDowell assures us, to “empirical realism” (not to be confused with a common description of one part of Kant’s doctrine) – a dualistic picture of the relationship of mind and nature in which contact is effected through, what in empiricist tradition is known commonly as, “the Given”. Ross (1970, 12) notes that, “The expression ‘the given’ became an accepted term of significance in contemporary philosophical vocabulary largely because of its use by C. I. Lewis in his *Mind and the World Order*.” He also notes (Ch. 2, §1) three ways of understanding the Given (as “qualia”, as the object, as such, and as a

⁶⁹ The “image” of disenchantment is originally due to Weber (McDowell, 1996, 70). The early sections of MacIntyre, 1987/1981, point also to the dominance of this image.

subliminal “blur” he attributes to Bradley (Cf. Taylor on Heidegger and Merleau Ponty (2002)). In McDowell himself, however, it seems that the idea of the Given does not (or need not) correspond to any of Ross’ definitions. In his 1998 lectures McDowell restates his debt to Wilfrid Sellars’ rejection of the Myth of the Given (Sellars, 1997/1956) and (while he accuses Sellars of not quite shaking off the myth) shares his specification of it. It is expressed in *Mind and World* in the following way:

Suppose we are tracing the ground, the justification, for a belief or a judgement. The idea is that when we have exhausted all the available moves within the space of concepts, all the available moves from one conceptually organized item to another, there is still one more step we can take: namely, pointing to something that is simply received in experience. It can only be pointing, because ex hypothesi this last move in a justification comes after we have exhausted the possibilities of tracing grounds from one conceptually organized, and so articulable, item to another. (McDowell, 1996, 6).⁷⁰:

Hence the Given is conceived as a non-conceptual warrant for what the thinker professes to be the case. Thus, in the empiricist epistemology McDowell thinks written into scientific methodology, what effects constraint on what is judged to be the case about the world is understood to be a non-conceptual deliverance, that, just as such, offers a ground for a claim to veridicality of an empirical proposition.

⁷⁰ Soffer (2003) states that, for Sellars, “the given is the immediate in the sense of the unlearned. It is the content of an awareness which does not require prior experience or training to be grasped but is simply “there,” in the raw. It is a type of awareness which does not presuppose language, concepts, or inferences.” However, whilst “rawness” communicates its uninterpreted character, to call the Given “a type of awareness”, on Sellars’ account of it, presumes, I think incorrectly, that it need have, for him, more than subliminal status.

The picture is that of a space in which conceptual activity takes place bounded by a non-conceptual sphere that is knowable via external impacts; the impacts, being non-conceptual, serving as warrant for the truth of claims made about that non-conceptual sphere. What worries McDowell about this traditional empiricist account is that if we suppose there to be an occurrent connection or contact between mind and a non-mental exterior we fail to accommodate what he calls, in his self-conscious echoing of Kant, *spontaneity*; the freedom we enjoy in judgement, in considering the correctness of our claims with reference not to a binding restriction on what can be correctly thought but to the rational support provided. On this empiricist model we are compelled to acknowledge the veracity of judgements that follow via an inexorable deductive logic from premises whose content is purely ostensive. We obtain an explanation of the source of a warranted belief that threatens to exclude willed avowal of its content. Or that, in McDowell's words, "offers exculpations where we wanted justifications." (1994, 8). On the other hand, if we remove the non-conceptual bridge between thought and its subject matter and suppose the primitive ingredient in our thoughts to be a phenomenal representation we are in danger of losing contact with an extraneous sphere altogether⁷¹.

⁷¹ The same worry is expressed by C. I. Lewis: "If nothing were given to the mind..then knowledge must be contentless and arbitrary, for there would be nothing of which it must be true. On the other hand, if there were no interpretation offered by the mind, then thought would become superfluous, the possibility of error inexplicable, and the distinction between true and false almost meaningless" Lewis' solution is, contra McDowell, to give a foundational role to the Given: "There are in our cognitive experience, two elements, the immediate data such as those of sense, which are presented or given to the mind, and a form, construction, or interpretation, which represents the

Thus, we are seemingly destined to oscillate in an attempt to accommodate both spontaneity and constraint in our thinking, caught, that is, between an empiricism that fully determines warranted judgement through a ground level input, and a coherentism that removes all external warrant; that commits us to a “frictionless spinning in a void.” (1994, 11). McDowell goes on, however, to note that this oscillation is inescapable only for those who feel the need to incorporate in their account of experience, and of knowledge, a bridge between what is conceptual and what is non-conceptual. But we need not, thinks McDowell, believe that the options have to be arranged around such a connection. That we do, in fact, feel this need is due, he says, to a fixation with the perspective, outlined above, that arrived with modern science. If we conceive of the relationship between mind and nature as one in which a sphere of conceptual activity is bounded by a realm extruded of thought entirely then a question is, of course, begged concerning how the two come into contact. The empiricist supplies a bridge through, in McDowell’s assessment of the position, a non-conceptual connection, while the coherentist rejects one half of the picture making the conceptual “unbounded”. But we can, McDowell thinks, conceive of the relationship between mind and nature in a way that avoids the seeming necessity to accommodate a duality of conceptual interpretation and at-point-of-entry uninterpreted input. His rejection of a view expressed by Davidson is instructive in bringing to light McDowell’s own, positive, thesis.

activity of thought." (C. I. Lewis (1929) "Mind and the World Order" (New York: Scribner), pp. 38ff, cited in J. J. Ross (1970).

In, for example, his *Empirical Content* (1986), Davidson defends a view that, on McDowell's reading, fails to provide a constraint on thinking from outside of the conceptual. Davidson's view is, McDowell believes, infected with the seeming obligatoriness of the scientific-empiricist distinction between thought as realm of conceptual activity and nature as exterior to thought. On McDowell's reading of him, Davidson recognises that a non-conceptual input cannot operate as ground for a justification for an empirical claim. However, he sees it as imperative for theory that the two realms be connected somehow. His approach is to provide a connecting informational bridge which is understood to have a causal but a non-justificatory role. That is, like McDowell, Davidson thinks that the intrusion of a non-conceptual input into a justification of what is thought reduces the thinker to passive commentator – reporting on what she is anyway bound to think, given the way rational moves in a justification link up ultimately to an in itself inarticulable presence. However, unlike McDowell, Davidson feels the need to accommodate somehow a connection that bridges, avoids a gap between, the conceptual and what is extra-conceptually the case. There is, however, in this account, McDowell argues, an unfilled gap between the non-conceptual ground that cannot, as such, operate as foundation in a justification and that which is understood to perform such a role. This unfilled lacuna leaves the thinker “spinning in a void”, justifying beliefs with other beliefs none of which connect to what is exterior.

To move towards an account of McDowell's own thesis we can note that he distinguishes between two senses of the term “content” (1994, 3-4). In one of its senses, the content of a thought is its subject matter. In this sense the content can be articulated, as what completes the phrase “the thought that.....”. Let this be

content(i). In a second sense content is the non-conceptual data understood as prior to and necessary for the formation of thoughts. In the first sense content is spoken of as “representational content” (or as propositional content). In the second, content is what has been referred to in empiricist tradition as “the Given” (call this content(g)). The latter, in the empiricist view McDowell attacks, supplies warrant for the veridicality of the former. In the light of this distinction we can see what is McDowell’s objection to Davidson here. In placing the Given, content(g), outside of a justification, representational content, what it is that is the content(i) of thought, is disconnected from an objective world which it putatively represents. Thus Davidson takes us into the coherentist side of the oscillation. McDowell’s own approach involves rubbing out content(g) and introducing a conception of the relationship between thought and nature in which we need not look for a ground for content(i) in the non-conceptual, as such. He rejects as unworkable the hypothesis that what warrants our beliefs is a non-conceptual given. The latter cannot, as Davidson recognises, function as both raw non-conceptual input and as justificatory ground. To play the latter role an interpretation is required of the purely non-conceptual that makes it impossible to conceive of it as such. And in an understanding of experience that recognises the incompatibility of the two roles, the Given drops out as an imaginary and otiose feature. Such a basis for a superior solution is, according to McDowell, obscured to Davidson because of the latter’s unexposed commitment to the “primitive metaphysic” (1994, 82; 1995, 164) implicit in modern scientific method. As has been mentioned, an unexposed “scientistic realism” (1995, 157) implicit in the modern outlook that conceives of an “in-itself” located beyond the conceptual sphere, imposes a requirement on

accounts of experience to supply a connection between the conceptual and non-conceptual.

In the view McDowell criticises, access to the non-conceptual is achieved through conceptual engagements at a basic level; a level of conceptual activity to be contrasted with the more sophisticated, higher level operations at which, say, law-like behaviours are considered. On this picture, that representations (content(i)) faithfully report or mirror an external world is certified by the transmission of uninterpreted data at the level at which non-conceptual and conceptual engage – this “empirical substance” being carried through “inferential linkages” to the level of higher-order concepts (1994, 6-7). McDowell’s move is to interpret the relationship between thought and nature in such a way that he avoids what he sees as the mistake of supposing the need for the Given in an account of experience. McDowell says that “experience is passive” (1994, 10). And it is very important to be clear in what way this is meant. As I read him, what McDowell intends is the thought that it is wholly beyond our will to alter the way in which the world *appears* to us. He goes on,

In experience one finds oneself saddled with content. One’s conceptual capacities have already been brought into play, in the content’s being available to one, before one has any choice in the matter. (1994, 10. The thought is repeated at p.31).

What the quote suggests is that the most primitive element in experience is to be understood as itself a conceptualisation. In the articulation of what is irreducible in experience what is articulated must be capable of being so. A pure sensory input is, as such, inarticulable and, being so, drops out of an account of what are

the grounds for the content (content(i)) of one's thought. In this way we can trace a ground without thinking that that ground outstrips the conceptual. The primitive conceptual activity associated with the receipt of information about the world is, it has been noted, passive. Thus, in the sense of passive in play here, the will is excluded from the ground level conceptualisations in which how things are is articulably yet unavoidably so. Being articulable items the most basic conceptualisations are suited to be integrated into a piece of reasoning that at a less basic level does involve the will. Our "engagement with" the world, one might say, is concept-involving but will-excluding, while, at a higher level of sophistication, thought is both concept-involving *and* will-involving. McDowell says:

"Thought" can mean the *act* of thinking; but it can also mean the *content* of a piece of thinking: what someone thinks. Now if we are to give due acknowledgement to the independence of reality, what we need is a constraint from outside *thinking* and *judging*, our exercises of spontaneity. The constraint does not need to be from outside *thinkable contents*. (1994, 28).

The constraint comes from outside *thinking*, but not from outside what is *thinkable*. (ibid.)

The way that we are aware of our environment is through unwilled (will-excluding) conceptualisations. Experience reduced to its primitive elements is not, then, according to McDowell, a series of subliminal sensory impacts. Experience is, for us, of what is expressibly the case; of, in the Wittgensteinian phrase McDowell is fond of using, what is "thus-and-so" (esp. McDowell, 1994, Lecture II).

§3.2 McDowell's Ethics

Equipped with the essentials of McDowell's account of the character of experience we can begin to explore how it informs his ethics. As with his epistemological concerns, central to McDowell's thoughts about ethics is the role a tacitly accepted conception of nature plays in existing metaethical positions. A leading source for this aspect of his moral theory is his *Two Sorts of Naturalism* (henceforth, *Two Sorts*), a paper which deals quite explicitly with this theme. McDowell notes there (McDowell, 1995, 149) that Philippa Foot as well as the "subjectivists" (recall "bald naturalists") and "supernatural rationalists" (recall "rampant Platonists") she makes her target are all committed to this restrictive conception.

In *Two Sorts*, McDowell offers an alternative to what he says is a common reading of Aristotle's ethics. On this reading of Aristotle's ethics the achievement of moral education is the exercise of restraint and willed direction of the passions according to principle. The mistake lies, so says McDowell, in supposing Aristotle intends us to imagine reason directing a distinct and fully-formed character whose propensities, and attachments are ready-made. On the reading he rejects, the rational appeal of ethical requirements and recommendations is grounded in evident facts about the relationship between humans and their environment. According to this account the virtuous life is the demonstrably right one to lead given a social-cum-natural context in which the claims of each are pitted against the claims of others under circumstances of scarcity. Further, the standard reading, so McDowell claims, appeals to the necessity of the relationship between virtue

and nature (drawing (1995, 149) on Aristotle, *Metaphysics* Δ. 5. 1015^a22-6). What this reading depends upon, so McDowell tells us, is an understanding of the good (or *Eudaimonia*, usually translated as “happiness”) as specifiable independently of the virtues. Only then can we read Aristotle as intending that we take the good to be achievable through exercising the virtues. However, McDowell goes on:

Aristotle explains ‘happiness’ as ‘acting well’ with ‘well’ glossed as ‘in accordance with the virtues’. The good he represents as unattainable without the virtues just is virtuous activity.” (McDowell, 1995: 150) (McDowell refers us to *Nicomachean Ethics* 1.7).

In response to the question, “Why be virtuous?” the response on the rejected reading is “Because it is necessary to achieve the good”. However, if the good just *is* virtuous activity the answer merely describes what will be, quite obviously, the state of affairs once the virtues are exercised. What is not supplied is reason to be virtuous, and the good understood as virtuous activity cannot provide a basis for a reply to that. A further related aspect of a common reading of Aristotle is rejected by McDowell, and consideration of this will further help us to understand his own position. An example of this disfavoured reading can be found, I think, in, Gaut (1997) in which the latter attributes to Aristotle a “recognitional model” of the role of practical reason in ethical behaviour:

This model holds that the goodness of actions and of states of affairs in general is constituted independently of those actions and states of affairs being the objects of rational choice. The role of practical reason is then to recognise the obtaining of goodness, and to bring it about that the agent performs good actions and brings about good states of affairs. What makes it rational to do such things is that they are good. (Gaut, 1997, 162).

Readings such as Gaut's present an unstable accommodation of two features of Aristotle's ethics: a naturalism that grounds the rational appeal of virtue in facts about us in relation to others and our common environment, and a characterisation of what virtue requires as being available only to those adequately sensitised (they being the practically wise). The recognitional view divides Aristotle's explanation of ethical behaviour into two parts. The first part supplies a foundation for virtuous behaviour in natural facts available to anyone. In Gaut's words,

Practical wisdom for Aristotle is a rational state, which grasps what is good or bad for humans....What is good is defined with reference to the human function, so that the good is constituted independently of being the object of rational choice. (ibid.).

The second part supposes that only those with the appropriate sensitivities can perceive what behaviour virtue requires. Gaut again:

the phronimos (the possessor of practical wisdom) has a sensibility conditioned by the virtues, which are exhibited in his knowledge of how to do the right things in the right way for the right reasons, and which require a grasp of particulars. (Gaut, 1997, 164).

On the one hand the phronimos (the possessor of practical wisdom) is one who is sensitive because virtuous and on the other hand one who is virtuous because he has recognised the good. But how can the good be recognised by one who is not antecedently appropriately sensitised? To think of the good as existing anyway, independently of the perceiver's sensitivities, is to view it from a perspective that is not available to the phronimos; indeed is not available to anyone not

appropriately sensitised. The reading has us reason in a circle: what virtue requires is not available to one who is not antecedently sensitised, while sensitivity develops from an apprehension of what virtue requires. The circle is analogous to the oscillation McDowell identifies in accounts of experience. What is required in both areas is a constraint on thinking that operates as an appeal to objective fact. Here, if we bring sensitivities into the picture as the source of moral understanding we threaten to remove objective ground. Once we put objective ground in place, in the manner of the recognitional model, however, we threaten to take moral understanding out of the hands of the individual. Accommodating both a role for sensitivity and for objective, interpersonal, ground is awkward for a conception of nature that places what one is sensitive to out there among the non-conceptual facts. The source of the misreading of Aristotle McDowell observes is again a conception of nature in which it is extruded of conceptual content. Here the crime is the ascription to Aristotle of an understanding of nature that post-dates his own time. McDowell leads us out of the circle with a re-conception of nature more akin, he thinks, to Aristotle's own. And central, here, is that for McDowell's Aristotle the value of virtue is intelligible only to those who already, perhaps unreflectively, practice it. Virtuous behaviour, as a category of action, is available as a category only to those for whom virtue is already of value: to those who know "the that", though not necessarily "the because". (Burnyeat, 1980, 71)⁷².

⁷² Here Burnyeat quotes Aristotle: "...while one must begin from what is familiar this may be taken in two ways: some things are familiar to us, others familiar without qualification. Presumably, then, what we should begin from is things familiar to us. This is the reason why one should have been brought up in good habits if one is going to listen

To engage in reflection about the nature of morality (to address “the because”) it is necessary that the individual recognise what acts, say, are kind, brave, tempered and so on. Such recognitional, cognitive capacities are what constitute “the that” and it is their acquisition that receives inadequate explanation in the reading McDowell rejects. And he joins Burnyeat in attributing a developmental model of morality to Aristotle. Burnyeat says that,

(Aristotle) knew intellectualism in the form of Socrates’ doctrine that virtue is knowledge. He reacted by emphasising the importance of beginnings and the gradual development of good habits of feeling. (ibid., 70).

And for what those “beginnings” are, Burnyeat again:

A wide range of desires and feelings are shaping patterns of motivation and response in a person well before he comes to a reasoned outlook on his life as a whole, and certainly before he integrates this reflective consciousness with his actual behaviour. (ibid.).

adequately to lectures about things noble and just..". Thus the appropriate student for Aristotle's series of lectures is one already schooled in moral requirements, and so suitably sensitised. A person not so equipped does not see the world in the required way and so cannot bring into view moral phenomena. Burnyeat clarifies: "The contrast here between having only 'the that' and having both 'the that' and 'the because' as well, is a contrast between the knowing or believing that something is so and understanding why it is so". And, also Burnyeat, 1980, 71: "The man who knows for himself is someone with 'the because' – in Aristotle's terms he is a man of practical wisdom equipped with the understanding to work out for himself what to do in the varied circumstances of his life – while the one who takes to heart sound advice learns 'the that' and becomes the sort of person who can profit from Aristotle's lectures."

This view, like the rejected reading, secures a place for natural facts about humans, our social and environmental circumstances, in an explanation of moral behaviour. Where they differ is that for Burnyeat and McDowell the way in which natural facts about us shape our behaviour is more subliminal than the recognitional model allows. That there are ethical requirements can be known only through a shaping of one's intellect and emotions in moral education. And, here, McDowell's philosophy of experience informs his ethics.

For McDowell, what action, say, is right in given circumstances is not something that should be understood as holding, or being the case, independently of the individual's interpretation of what is, in that situation, required. This basing of a correct judgement on the sensitivities and beliefs of the individual without McDowell's revised conception of the relationship between nature and mind will lead some, of course, to doubt that he has secured objectivity. This will only be the case, however, if we think of objective warrant as available in a way that exceeds what a mature adult can possibly come to think of as morally required. In this way, as with his epistemology, an external realm outside of how we can possibly conceive the world, drops out of view. In the moral realm what is perceived as required is a function of upbringing and determines in part what is seen as required in a given situation. If we absent from the history of an individual an upbringing (or other means) in which an appropriate refinement of the "moral sense" is achieved then we remove the possibility for that individual of seeing things in the way a morality requires. How the world is for us is limited to what can possibly be the content of a conceptualisation; and this in a way that removes the mistaken search for a non-conceptual source of what is thought. This permits a

conception of moral perception that sees it not as a spreading of value or emotion on a factual canvas but as the isolating of what is an integral element in the apprehension of a scene.

What is of import for us is not the contribution this makes to a discussion about moral realism but the capacity it imbues the agent with to reason toward action from premises that invoke the contents of a moral outlook. It is an important feature of McDowell's ethics that we be conceived not simply as applying rules and responding to ethical signals we have learnt to identify, but that a moral outlook comes replete with a capacity to assess its validity. And this is a rational capacity in which what our desires are, facts about what is possible, moral principles and so on are all in view as contributing to an assessment of what should be done:

Reason does not just open our eyes to our nature, as members of the animal species we belong to; it also enables and even obliges us to step back from it, in a way that puts its bearing on our practical problems into question.(McDowell, 1995: 154).

I do not mean to suggest that McDowell's approach solves the problem of how to support a conception of reflective thinking that combines both embeddedness and critique. He recognises that the Aristotelian model he draws on pays insufficient attention to self-examination, the shaping role of the *polis* being emphasised over the need for watchfulness at the individual level. And McDowell's own adoption of a Neurathian conception of reflective critique will not silence critics who, like Habermas, seek a source of universal principle that transcends tradition. However, on an, I think, uncontroversial reading, McDowell offers a picture of

evaluative deliberation in which there is a place not only for rational capacities that enable an individual to think within a tradition, but also for neutral critical faculties that enable a self-awareness that sees the tradition-bound modes of thinking as inherently open to critique. This is not to say that the person can step outside of a tradition and assess its merits wholesale. What it does point to is the special way in which commitments, and the like, are *held*. While neutral rational capacities may not permit extra-tradition inspection of standing commitments, and the like, they do affect their status. They are, on McDowell's model, inherently open to critique, a characteristic that forms the basis for a middle way between a ("bald naturalist") understanding of beliefs and so on that rubs out their vital background support, and a (rampant Platonist) conception that invests them with directive status within a living morality.

§4. A McDowellian Reading of Rawls

On McDowell's Neurathian conception of reflective thinking it is necessary that the reasons, beliefs, modes of evaluation, and so on, that inform a moral outlook are held in a way that sees them as inherently sensitive to critical inspection, and, indeed, revisable on a piecemeal basis. Moreover, the reasons, affections and so on, that are viewed as so sensitive also retain, on this picture, their practical purchase, while they are the subject of criticism. In McDowell one finds oneself landed with a living morality, equipped, nonetheless, with the means to assess its justificatory basis. Such a conception can, I want to show, be identified in Rawls.

In lecture II, section 7, of *Political Liberalism* Rawls attempts to aid our understanding of how he conceives the evaluative thinking of his citizens, by

characterising types of desire they are presumed to have. Their desires can, he says, be categorised as object, principle, and concept dependent. First, object-dependent desires. These include desires for food, drink, social status, wealth, and desires corresponding to attachments, affections and loyalties (the last three might be exemplified, presumably, by a desire to see a friend, or one's national sports team do well) (Rawls, 1996, 82). What marks out a desire as object dependent is that,

the object of desire, or the state of affairs that fulfills it, can be described without the use of any moral conceptions, or reasonable or rational principles. (ibid.).

To account for a football fan's applause for her team it is adequate to mention her desire to see them do well, which desire can be accounted for by a description of her special attachment to that team. This pointing to a desire ought not, however, be taken to imply a Humean picture, for this type of desire. On the (canonical) reading of Hume desire is the primitive spur to a motivationally inert reason, and desire does not have this foundational role in Rawls' theory. To see why, or in what way this is the case, it is helpful to consider what he says about principle-dependent desires. He says,

the object or aim of the (principle-dependent) desire, or the activity in which we desire to engage, cannot be described without using the principles, rational or reasonable as the case may be, that enter into specifying that activity. (ibid.).

Examples Rawls gives, of principles that serve to make principle-dependent desires, and associated activities, intelligible, are the following two rational

principles: to adopt the most efficient means to given ends, and to select the “more probable alternative”. Examples of reasonable principles include those “associated with the moral virtues recognized by common sense such as truthfulness and fidelity.” (ibid., 83). Hence, if we envisage an actual choice of some item of food, say, we can, or it is explanatorily adequate to, account for the action by mentioning an extant desire quite simply for food, or that specific item. An act of truthfulness, on the other hand, will not, in a normal case, be attended by a desire fully described as one to be truthful just on that particular occasion, but will reflect a principled commitment to honesty. And, in order to make this sort of case different to the object-dependent case, the commitment itself must not be taken to have a desiderative ground. Were it to be so taken then the principle itself could be the object of the desire, and the principle-dependency would in fact be mere object dependency. Thus, the object and principle dependency of desires can be distinguished by pointing to the need to posit commitment to a principle in order to rationalise certain types of behaviour (eg. Promise keeping) and not others (eg. Buying a personal stereo). But there is more to the distinction. Rawls intends to characterise his agents’ evaluative thinking in such a way that they possess “autonomy”: understood as a capacity to exercise self-command, to act from principle in a way that denies a heteronomous role to affection.

In a footnote, that is very significant in this context, he says:

It is important to stress that the force, or weight, of principle-dependent desires is given entirely by the principle to which the desire is attached, and not by the psychological strength of the desire itself... A person with a good will, to use Kant’s term, is someone whose principle-dependent desires have strengths in complete accordance with the force,

or priority of the principles to which they are attached. (ibid., 82-83, n. 31. My emphasis.).

Hence, there is an intellectual appreciation of the rational weight of a principle, where the rational weight, or its intellectual appeal, itself has a motivational role, quite independent of a, factually present, desire to act on that principle. Indeed, I think, a correct interpretation of Rawls, here, sees the factually present desire as dropping out of Rawls' account, being replaced by desire understood as that which accounts, as a matter of description, for the motivational attachment to a principle, or, indeed, an object, but that does not do the generative role of desire in a standard Humean account. In support of this, note that Rawls applauds Bernard Williams' work here, on the "motivational set", a conception, Rawls tells us, that may dissolve the traditional divide between Humean and Kantian accounts of practical deliberation (ibid., pp. 84-85 & 85, n.33).⁷³ Autonomy, non-desiderative self-direction, may, if Rawls' account is correct, be possible without the metaphysics of a noumenal will. Indeed, what Rawls is effecting here, he seems to say, is a theoretically plausible account of practical deliberation that captures our intuitions about what evaluative reasoning must be like for there to be social agents capable of reasoning about justice. (the model he presents draws importantly also on Thomas Nagel (1970)). This is not to say that an object-dependent desire cannot compel. It might be a fact about an agent that he simply cannot resist being drawn towards food. Rawls' model is, I take it, consistent with akrasia as it is commonly experienced. However, for object-dependent desires to fit into the overall scheme that Rawls presents it cannot be the case, I suggest, that

⁷³ See §2 above.

object-dependent desires compel just as natural facts about a person. It must be the case, rather, that they are responsive to reason, and this nullifies their status just as natural facts, simpliciter (what Taylor would call “de facto” desires: see §2). To see why, we need only recall from what was said above that a principle-dependent desire can be characterised as merely object-dependent. A persons’ sustained fanaticism for a certain principle might make his attachment to it pathological. The object in this object-dependent desire just is the principle. Suppose, then, as is natural, that we want to offer an account, using Rawls’ scheme, for, on the one hand, a person whose formerly rational commitment has become pathological, and, on the other hand, the remediable nature of his condition. In either case, were we to think of his (pathological) object-dependent desire as a natural fact about him, simpliciter, as in a basic Humean account, then we would not have the means to offer a narrative that involved his desire, as in the second case, for example, coming to be, once more, principle-dependent. There would in our story be a transition between desires of distinct status (from object-dependent desire to principle-dependent desire) with no intermediate status linking the two. The way to remedy this is to think of his object-dependent desire as already, so to speak, saturated with reason, and so responsive to rational control, in the way that a desire, just as natural fact, is not.

But there is more to the inherent responsiveness to reason of desire, more broadly, affection, simpliciter. A conception of desire that sees it as responsive to rational critique need do no damage to a simultaneous understanding of it as psychologically directing. While this latter aspect is not, one suspects, embraceable in thought, a conception of it that preserves its directive character

can form part of a rational understanding of it. The notion of conception-dependent desires completes the Rawlsian picture of self-regulation. It would be possible to imagine a person whose commitment to principles is unreflective, the non-desiderative self-command he exercises being based in an unreflective commitment to social norms (the above pathological example is a cognate, though dissimilar, case). It is possible, from a third-person perspective, further, to imagine a person selecting in a norm governed way among competing principles. This would not require the person to have a reflective grasp of the reasoning that supports the selection of possible principles. But this is not what Rawls wants. Conception-dependency refers to the selection of principles according to a rationale that reflects a *knowing commitment* to substantive ideals:

These desires can be described by saying that the principles we desire to act from are seen as belonging to, and helping to articulate, a certain rational or reasonable conception, or a political ideal. (ibid., 84).

Rawls doesn't fully spell out what he intends by his use of the term "desire" in this context. It is to be expected that he intends us to think there is a difference in type between a desire to eat fruit and a desire to apply a principle, but Rawls doesn't expand. I think, however, that, insofar as what Rawls has to say bears the interpretation I have offered, there is, at least in one sense, parity between object, principle, and conception dependent desires. If an object-dependent desire is not a desire simpliciter then the conception of it will be integrated with its rationale. A person's desire to picnic with friends cannot, on this understanding, be reduced to descriptions drawn from human physiology but must mention phenomenal aspects

of, for example, the nature of his relationships. And in supplying this sort of information we are at the same time supplying a reason for his desire, in a way that has the desire and reason appear as a couplet. Indeed, there is no way to the desire apart from through its attendant rationale. Similarly, and more patently, a desire to see a principle inform one's action will be inexplicable just as such, but must mention the associated rationale. But then what emerges is not only the autonomy of evaluative reasoning from desire as such, but the independent status of reasons. That is, they are responsive to critique but are also simultaneously part of a living morality, hence capable of playing a shaping role in the person's deliberations.

Note, with this, the following:

while a reasonable comprehensive view is not necessarily fixed and unchanging, it normally belongs to, or draws upon, a tradition of thought and doctrine. Although stable over time, and not subject to sudden and unexplained changes, it tends to evolve slowly in the light of what, from its point of view, it sees as good and sufficient reasons. (Rawls, 1996, 59);

We have a picture of normative thought, then, in which reasons with independent status, in the above sense, are informed by a tradition, that, whilst encumbering, does not foreclose critique. As possessors of a comprehensive doctrine, then, Rawls' agents normative outlooks bear a Neurathian interpretation. Rawls can, then, be read in such a way that he is not guilty of requiring a detached perspective that nullifies the generative factors in a thick, ethical vision. But what can be said of the relationship between the citizens qua citizens, replete with comprehensive doctrines, and their specification in the original position? Even if

Rawls can be said to present a conception of the person that facilitates incremental change to a comprehensive doctrine, through its openness to living critique, it might be the case, of course, that the original position (he invites his citizens to consider, as a device for the derivation of public principles) requires they neutralise the bases of their thick perspectives.

In his reply to Habermas (1996, Lecture IX) Rawls appears to close the option of supposing political discussion to somehow integrate with that associated with a comprehensive vision: "...political liberalism moves within the category of the political...It leaves untouched all kinds of doctrine – religious, metaphysical, and moral – with their long traditions of development and interpretation." (ibid., 375). However, it is, as ever, important to remember the different standpoints in Rawls' theory. This statement (he calls it a "central idea" (ibid.)) is made from the point of view of the theorist and so need not be taken to imply a similar bifurcation in the minds of the citizens themselves. The citizen need not adopt the theorist's standpoint in assessing the rationale for the two principles of justice and so a strict compartmentalisation (of political and comprehensive discourse) may not be expected of him. Note first (p. 374) that for a person to be "reasonable", that characteristic of a person that enables them to compromise, they must possess,

a willingness to propose fair terms of social cooperation that others as free and equal also might endorse – and to act on these terms, provided others do, even contrary to one's own interest; and second, a recognition of the burdens of judgement and acceptance of their consequences for one's attitude (including toleration) toward other comprehensive doctrines.

Now, Rawls discusses the “burdens of judgement” (Lecture II, §2) in the context of explaining how “reasonable disagreement” is possible. Here he refers to the two moral powers and says (p. 55) that,

Given their moral powers, they share a common human reason, similar powers of thought and judgement: they can draw inferences, weigh evidence, and balance competing considerations.

So that what makes a comprehensive doctrine reasonable is the capacity on the part of the individual whose doctrine it is to engage in balanced discussion with other doctrines. The “burdens of judgement” are then the capacities for neutral (or, perhaps, inter-doctrinal) thinking that accompany a “reasonable” comprehensive position. An alternative way of specifying the “burdens of judgement”, one that brings out more clearly the idea of burden, is to see reasonableness as imposing requirements for impartiality, and the like, on the individual. This bears a striking resemblance to McDowell’s thought that the “freedom that...is...exemplified in responsible acts of judging, is essentially a matter of being answerable to criticism in the light of reasonable considerations.” (McDowell, 1998c, 434). There is a difficulty here associated with the type of access individuals can have to other doctrinal outlooks, but Rawls need not be taken to be specifying an out-of-doctrine point of view, only a degree of openness that permits cross-doctrine discussion. Indeed, Rawls is not looking for agreement. What he seeks is a basis for reasonable disagreement, and this does not require either an out-of-doctrine purview or first-personal insight into others’ way of seeing. So, we begin to see the seamlessness in, what we can call, Rawls’

conception of “embedded critique”. Though shaped by different comprehensive doctrines, persons’ also possess, in common, the rational capacities to work towards and sustain a material, if not an ideological, compromise. But it might still be remarked that, even if it is accepted that the citizens are represented as embedded in an ongoing public culture, and, further, that they are capable of inter-doctrinal debate without stepping outside of their comprehensive outlook, it is still the case that in the original position comprehensive doctrines are blind-sided. And this makes inoperative what are, in fact, directive in living deliberation – so making for an unrealistic account of the bases of social stability. One response to this goes as follows. In *A Theory of Justice* Rawls views his approach as a specific application of the theory of rational choice. But later Rawls disavows this remark (1996, 306, n.21). He says in his later work that,

(The original position) helps us to work out what we now think, once we are able to take a clear and uncluttered view of what justice requires when society is conceived as a scheme of cooperation between free and equal citizens from one generation to the next. The original position serves as a mediating idea by which all our considered convictions, whatever their level of generality – whether they concern fair conditions for situating the parties or reasonable constraints on reasons, or first principles and precepts, or judgements about particular institutions and actions – can be brought to bear on one another. This enables us to establish greater coherence among all our judgements; and with this deeper self-understanding we can attain wider agreement among one another. (1996, 26, emphases added).

The reference to society as “a scheme of cooperation...from one generation to the next” in this quote is frequently made by Rawls, and allows us, quite confidently, to situate his citizens in society as an ongoing, organic enterprise. Stability,

indeed, is achieved, in part, not through a periodic thought experiment in which the principles that would be ground out in the original position are held against political reality, but through the strength of the public culture the principles support. It is from within an ongoing society that the citizens consider the justness of their living institutions – they do not enter into the society having first agreed to a way to design public rules. The rational capacities associated with reasonableness need not be thought of as separate from the capacities that are drawn on in an evaluation that calls into play the ingredients of a comprehensive doctrine. The tools of critique necessary to material compromise, to construct a public space in which competing ideologies can influence the activities of different groups, need not be thought of as extra to the capacities drawn on in carrying out those activities. We can think of the comprehensive doctrines as coming replete with the means to assess them, and the means to establish what, given the reasonable claims of others, it is reasonable to claim for oneself.

§5. Beyond the Market Context

Market-oriented choice, though a central aspect of everyday life in a modern economy, covers only a limited domain of personal choice. It leaves out choice, and, more broadly, conduct, in family and other social contexts, and so the picture of embedded critique I have argued for might be thought, for that reason, to be of limited interest to Rawlsian criticism. In this section I want to argue that the conception of embedded critique associated with market participation has implications for how we view the critical faculties of the Rawlsian citizen as such, and not only in his role as market participant. This can be done, I want to argue

(in §5.1), by drawing on the notion of metapreference I used in chapter 3 to formulate a conception of embedded critique. This can be fleshed out, however, with reference to the picture of ethical evaluation I have found in McDowell's moral philosophy. Rawls himself, in *A Theory of Justice* if not in *Political Liberalism*, takes two approaches to the characterisation of his agents - approaches that, when combined, present, so he intends, a complete picture. In the third part of the earlier work Rawls talks about the moral development of persons, their upbringing under the authority of adults responsible for their care, and so on, presenting an empirical, developmental account of moral personality. This contrasts with the presentation in part one of *A Theory of Justice* in which the analysis proceeds with reference to the rational (and reasonable) capacities of the individual considered quite apart from the realities and contingencies of moral upbringing. I presume from this that it was part of Rawls' intention to provide a conception of the person that accounted both for the rational capacities necessary to neutral, moral evaluation and for the way in which what is morally possible for persons is determined in large part by what can be transmitted across generations, and the ways in which family and community can shape moral character.

The way in which I defend an extension of my thesis to the Rawlsian agent per se is intended to mirror this type of approach. I begin with an analysis that refers only to evaluative criteria and their rational support, and go on, drawing on McDowell, to refer to the contingencies of moral upbringing. In such a way, I hope, that there is a neat fit between the conception of reflective critique associated with the market participant and that associated with the more developmental picture I draw from McDowell. My intention is that the way these two fit together

preserves the rounded conception Rawls, I have suggested, wanted for his theory while avoiding the criticism that his representation of the stable society is unrealistic.

§5.1 Metapreference and the Basic/Non-Basic Distinction

One, perhaps, fruitful way to argue for a more general view is to note the overlap between market institutions and those associated with other areas of life. Very often it is the case that a restriction in the marketplace reflects discussions not directed in the first instance at market conduct. Prohibitions (on what can be traded) intended to preserve public morals, for example, though they have implications for market choice will commonly originate outside of that context, and have implications for what counts as permissible conduct more generally. The institutional rules that the market participant confronts, then, will see him engaged in debates concerning public policy at large. Hence, the critical capacities associated with market participation should not be viewed as those of an isolated actor, isolated in that his critical response to other types of institutional constraints is very different. It is possible to view the market (as Karl Polanyi has classically argued (Polanyi (2001/1944)) as embedded in political and social structures. The market participant is, in this way of seeing things, caught up in a nexus of overlapping institutions that make it difficult to see the criticality associated with market participation as limited to the market domain.

However, among the objections that might be raised to the view that a general thesis is possible, is one that observes the lack of criticality among those who uncritically accept the constraints imposed by non-basic institutions, and

whose actions in the marketplace are, for that reason, without the supposed critical aspect. A devout Jainist, for example, will, as part of the vow of non-violence, be vegetarian, and will exhibit then, so the argument might go, uncritical choice behaviour in the market. He may choose, in this setting, to blind-side all possible evaluative criteria other than those that accord with his religious convictions. He may, for that matter, be inured in his Jainism so that his beliefs inform his choices as a matter of blind habit. This, surely, is a market participant who lacks the critical capacities I have argued for in connection with the market actor.

It will be helpful, for coherence, to put this contention in a metapreference framework. I argued earlier that the reflective capacities of the market agent are well represented by a metapreference model of practical reasoning. Here we can suppose the binary choice set, $C\{x, y\}$, to contain elements, x and y , that correspond to items proscribed and not proscribed by a person's religion. Element x might be meat, for example, while y is cereal. A person might, applying the metapreference model, have preferences $xP1y$ or $yP2x$, there being associated with each possible preference ordering an evaluative criterion that informs that preference. Reason-based autonomy entails that the alternative criteria be assessed from a point of view that brings both criteria into play, this point of view being one that cannot, then, be itself regulated, exclusively, by either evaluative criterion. Where E, E' represent alternative evaluative criteria we can (recalling chapter 3) adopt the following schematic representation of choice: if E then $(xP1y)P(yP2x)$ and $C\{x, y\} = x$, where the choice of E reflects an all things considered assessment of the reasons supporting E . In this way of presenting things someone who objects that I cannot go from a conception of critical

evaluation for the marketplace to a more general account could say that for the person habituated in the religious convictions of her community the choice of evaluative criteria will not be a matter of reflective decision. Where evaluative criterion E reflects her habitual religious convictions there will be no question whether E informs her choice, and so an alternative criterion (E') and, moreover, the reasoning that compares E and E' will form no part of her reasoning towards choice. Where alternative criteria are those associated with the institutional restrictions of a well-ordered Rawlsian society, those criteria, and their supporting rationale, will be similarly silenced. The person whose habitual religious convictions inform her market choices, then, will be a market actor, will be someone who can make market-based choices, but she will not possess the reason-based autonomy on which I have based my argument for embedded criticality. My argument, then, it could be argued, does not generalise. Indeed it rests on a spurious assumption of reason-based autonomy that it is improper to associate with market participants once we recognise the full extent to which market choices can be informed by personal convictions.

I would argue, however, that such a person has no place in a Rawlsian society, and that a strong support for my argument suggests also that it would be untenable to suppose that my thesis cannot be generalised to cover the Rawlsian citizen per se. To see this, imagine an agent whose market-based choices are exclusively informed by his comprehensive doctrine. Any evaluative criterion E that informs choice is supported by reasons that have a basis in his comprehensive view. This means that (staying with the metapreference conception) that he is capable of conditional judgements of the form, if E_i then xP_iy , but any E_i, E_j that

are compared will be have supporting reasons that are drawn from comprehensive doctrine. Letting superscript c denote that the evaluative criteria are supported by reasons drawn from comprehensive doctrine, we can not that any E_i , E_j should properly be designated E_i^c , E_j^c . What such a person cannot do, by assumption, is to bring into play in her reasoning evaluative criteria supported by reasons whose support is neutral with respect to comprehensive doctrines. That is judgements of the form E_i^c or E_i^n (where superscript n denotes neutral support) are not available. A Rawlsian agent, however, a citizen of a well-ordered society, as Rawls describes it, regulated by the two principles of justice, must be capable of making judgements of the form E_i^c or E_i^n . Were she incapable of such judgements then she would be incapable of thinking that did not engage her comprehensive doctrine, and so incapable of the type of thinking required for neutral assessment of her society's institutions. So the person whose choices are informed exclusively by habitual religious beliefs cannot be a Rawlsian citizen, and so, therefore, cannot be a Rawlsian market participant. But this line of argument supports the view that the Rawlsian agent qua market participant *and* qua citizen must possess the same hierarchical reflective capacities, capacities that enable the reflective assessment of evaluative criteria whose rational support stem from alternative (neutral and doctrinal) sources.

§5.2 McDowell and the Market Agent

What McDowell offers, and what was lacking in both Williams and Taylor, is a way to remove the *backdrop*, the settled modes of thinking, evaluative stances, and so on, most familiar, perhaps, from Williams' work, that regulates what a

person thinks about a policy, event, state of affairs, and so on. Being unable to get behind that backdrop we are unable to distinguish between the mode of reasoning a person adopts in her everyday life, and the mode she adopts in neutral scrutiny of public institutions. Rawls commends Williams' approach. And the latter does indeed offer a way to accommodate personal concern in a way that avoids undersocialisation. However, what Rawls requires for a realistic representation of social stability is not only the retention of the practical significance of personal concern in reflective deliberation. In order not to fall into comprehensive liberalism, a liberalism that presumes a common stock of comprehensive convictions, Rawls must also incorporate in his conception a capacity for out-of-outlook neutrality. Williams does not, I have argued, offer a position that can satisfy both desiderata. He fails, I would suggest, to offer room for evaluation the bases of which do not lie within comprehensive doctrine. Rawls' thesis, however, can be read in such a way, or so my previous section suggests, that the higher-order criticality, criticality that transcends substantive modes of thinking, I have associated with McDowell's work, can be associated with Rawls' work too. This is helpful for two reasons. First, the picture of reflective criticality it recommends lends support to the conception supported in chapter three by discussion of what is involved in market participation. Second (what I want to pursue here) it helps to establish further the implications for the conception of Rawls' agent per se, that is it helps us to take the argument beyond the market context.

McDowell, on the interpretation I have offered, identifies two ways of approaching an assessment of normative positions, ways that vie in moral philosophical tradition. We can adopt a naturalistic stance, that views normative

concepts as adequately described by the sort of language appropriate also to natural (and much social) scientific theory. To talk, for example, of the evolutionary role of moral rules or commitments, is to effect such a reduction, removing from view, as theoretically unimportant, the living significance of rules and commitments. An alternative approach emphasises the regulatory character of moral concepts. On this (recalling McDowell's terminology, "Rampant Platonist") view, moral concepts and categories are what do the work in moral thinking and are themselves, as such, not open to reduction. What makes these options seem the only options available is, McDowell says, (and I am suggesting that this is a good account of the way the analogy works between his epistemology and his moral philosophy) is a presumption that the facts (first-order) moral thinking takes into account must be either in themselves without evaluative significance, just raw facts, or must be irreducibly value-laden, and so without separable "raw" content. Given this understanding of the possible ways in which fact can appear or operate in moral thinking, if we want to avoid the sin of Rampant Platonism, the sin of investing the world as such with moral significance, then we can run into the arms of the naturalist. If we do, however, we remove from view the living, normative significance that makes the world we inhabit a moral world at all. The way out of this is to reject the seeming need to respect a view of nature that sees it as value-neutral. If we feel the need to respect this conception of nature we, of course, in Humean fashion, must see the source of value as the subjective feelings, emotions, and so on, of the spectator. Recalling the analogous manoeuvre in his epistemology, McDowell's step here is to reject the view (the myth of the Given) that nature must operate as conceptually undilute in a piece of (in this context)

normative thinking. The defect of the Rampant Platonist position is that it sits most happily with a view of morality in which the latter stains the world in a way that reflects its background presuppositions. If we no longer acknowledge that this must be the way in which we see nature as coming to having normative meaning, then we can view things at a more general level. We can talk, in a theoretically respectable way, about the potential meaning that facts might have for a morality given its commitments, and, moreover, we can make room in an actual morality for the possibility that the moral significance things have reflects faulty commitments, or ways of evaluation. This creates room, that is, for a conception of the role of criticality in moral thinking that does not require the living significance things have for its participants to be neutralised in an assessment of their merits. This possibility of a Neurathian movement in moral thinking, in turn, requires that we do not view a morality as possessing a regulative backdrop, but, rather, as characterised by a higher-order criticality not itself informed by background presuppositions. And it is this higher-order criticality that I have argued ought to be associated with the market participant, in the exercise of her reason-based discretion.

It is worth noting also the important contribution that is made to McDowell's thesis by the notion of *Bildung*. The term has a long philosophical history, and is associated with a rich range of definitions, and its use in McDowell is intended to convey, no doubt, some of this subtlety. However, *Bildung*, in a perhaps too prosaic usage, can be taken to refer to the moral upbringing of a child, where what the use of the term *Bildung* emphasises is the rational autonomy that is thereby acquired. On this view moral education develops a critical take on a

morality *pari passu* with knowledge and an appreciation of its substance. Now, evidently, people do not arrive as fully formed adults in the marketplace, knowledgeable of much of a society's moral culture but needing schooling in the ways of market trade. Part of our *Bildung* is an orientation towards the market, and, it is plausible to suggest, preparation for the exercise of reason-based autonomy in the marketplace constitutes a significant portion of the "moral" upbringing of the child in a modern market economy setting. So, there is, if my reading of Rawls in section four can be made to carry water, a rather satisfying connection to be had between the Neurathian conception of the person that can be found there and the conception of reflective thinking I have argued for in connection with the market participant.

§5.3 Beyond the Market

What is understandably missing from McDowell's account of moral education is the distinction, important in Rawls, between basic and non-basic institutions. People, of course, in actual as well as in well-ordered societies, are raised in the ways of non-basic institutions as well as being exposed to institutions that are, in the Rawlsian sense, basic. The explicit and implicit moral codes imposed within families, among friends, in professional contexts, and within religious faiths, exemplify institutional contexts that are non-basic. The market and, more generally, social and legal institutions are, in one sense, basic because their rules and regulations play an important role in defining non-basic institutions, and because their provisions, in part, govern relations among non-basic institutions. In this way what legally constitutes a family is governed at a basic level, while, in

some cases, what a person is permitted to do within the family (children's bed times, and so on) is non-basic. The property rights of religious communities are, for further example, basic, while their credos are non-basic. Now a person will, in his or her upbringing, be sensitized to, or be made aware of the need to be sensitive to, the basic requirements of market trade. These will include prohibitions whose justification, being basic,⁷⁴ does not draw on the comprehensive doctrine of any particular social grouping. A person will, typically, however, also be exposed to the prohibitions and prescriptions associated with the non-basic institution (indeed, institutions) she belongs to. It is quite possible that there will be no interference between basic and non-basic requirements at a personal level. While trade in dead animals is permitted at a basic level it might not be permitted at a non-basic level, for example. The individuals this affects can recognise both a basic permission and a non-basic obligation, and can act accordingly. Things might not be so simple for those (for example, fox-hunters) who want to engage in something impermissible at a basic level but permissible at a non-basic level. However, even here, there is no intellectual, or logical, difficulty involved in perceiving an obligation (a basic obligation) and a (non-basic) permission. Suppose, then, somebody whose *Bildung* has been comprised of exclusive inculcation in a basic institutional setting, so that she is blind to obligations and permissions other than those that emanate from her non-basic institution. Suppose, further, that one of her permissions conflicts with an obligation associated with a basic institution,

⁷⁴ The sense in play of "basic" here entails, of course, that a justification is basic if it is associated with a basic institution (here the market).

specifically the market. It seems straightforward that if *she* (and not merely her teachers) is to be a market participant then she must be sensitive to normative considerations of a basic type, and not merely those that are non-basic. Note further that she will have to be able to bring into view justifications associated with both basic and with non-basic normative considerations if she is to be sensitive to restrictions and permissions emanating from both quarters, restrictions and permissions that may concern the same objects of choice. Hence, so long as we cannot envision agents who are sensitive to only non-basic norms (and one good reason for not envisioning such people is that they will not qualify as reasonable in the required sense) we must have agents whose criticality embraces market-based norms. Once we have in place only such people then it seems untenable to suppose agents in family, and other non-market contexts, to possess the bipartite critical faculties the neo-Hegelian finds so objectionable in Rawls, while, qua market actor, her criticality transcends the neutral/non-neutral divide. It seems defensible, then, to view the critical capacities I have argued for in connection with market participation, as those of the Rawlsian agent as such, and not only in her guise as market agent.

§6 McDowell, the Market, and Stability

I attempt in this section to pull together the threads of the argument presented in this chapter, and to bring out their relationship to my overall thesis. I begin by recalling that Rawls devotes a substantial part of *A Theory of Justice* to the problem of stability. His intention, in Part III, appears to be to present a theory of moral development that shows how affective attachment to the basic institutions

of a well-ordered society is compatible with a belief that the behind-the-veil reasoning is an appropriate rendition of the moral point of view, for purposes of assessing existing (basic) institutions. The way he hopes to achieve this combination is by pointing to the universality of the interest in fair and stable institutions while emphasising the progressively greater rational distance a person is able to achieve from his desires, beliefs, and attachments as he matures.⁷⁵ Rawls comes to doubt the validity of such a combination as relying on universal interests presupposes sufficient resources for political stability in the agents' comprehensive doctrines. This, he comes to think, presupposes too much, and that the real challenge for a theory of justice for a democratic polity is to accommodate agents whose comprehensive doctrines are distinct in their fundamentals, so that compliance, hence stability, cannot be presumed to flow from shared comprehensive convictions. Rawls adjusted his theory to take these considerations into account and presented a conception of normative reflection for the public sphere (in *Political Liberalism*) that views agents as capable of evaluative thinking, with respect to basic political, economic, and social institutions, that does not bring into play their comprehensive doctrines. This bipartite construction, I have argued, generates an oversocialised conception of the obedient citizen that leads to an unrealistic representation of a stable society. Agents are construed as acting on themselves from outside of their comprehensive doctrines in order to override their personal motivations. They ought to be represented as vanquishing, or moderating the motive power of their personal concerns. Rawls' representation

⁷⁵ Feminist critics of Rawls would say that I am right to say "he" here, as Rawls' (Kohlbergian) account emphasises the development of what are predominantly male characteristics.

has them bypass their motivations and so fails to represent the tension that ought to be represented in a stable, democratic society between willed compliance and the vital, though restrained, sources of dissent.

The problem, then, of how to accommodate the commitment required for social stability and the rational detachment required for neutral thinking does not go away for Rawls with the adjustments he makes in *Political Liberalism*. The problem can be seen, not too stylistically, as how to accommodate in one theoretical framework an *embedded* commitment to basic institutions, embedded in the sense that the commitment emanates from agents' comprehensive doctrines, and a supporting rationale for those institutions that appeals to a moral point of view that is neutral with respect to comprehensive, doctrinal positions. Seen in this light, what Rawls requires is a conception of *embedded critique*, a conception that accounts for the commitment that flows from comprehensive doctrine and that allows sufficient reflective distance to allow neutral, non-doctrinal, discussion of shared, basic institutions. One way to construct such a conception is, I have argued, to draw on a characterisation of market participation, and such was the argument of chapter three. This allows room for commitment that emanates from comprehensive doctrine to be combined with rational reflection that subjects such commitments to rational critique. What is not supplied thereby is a complementary picture of moral thinking and its development that indicates how flesh and blood citizens in a well-ordered society can, in a way that is compatible with this (market-derived) conception, become thoroughly committed to, though potentially critical of their society's institutions. The moral philosophy of John McDowell was brought on board to provide that second part of the conception,

achieving, it is intended, the combination of commitment and stability, on the one hand, and moral neutrality, on the other, that I have argued is missing in both Rawls' earlier and later work.

§7. Conclusion

Rawls presents a positive view of theories of reflective critique that emphasise the embeddedness of personal criticism. He thinks that Bernard Williams, for example, has valuable things to say about the inability of the person to escape into rational detachment when assessing her moral and political worlds. However, thinks Rawls, this ought not to prevent us from describing an artificial conception of reflective criticism that has the agent capable of radical detachment. I would suggest that this combination of views is not open to Rawls because he must offer a conception of reflective criticism, in a well-ordered society, that not only maintains the vitality of personal concerns, something Williams' type of approach can secure, but also the type of neutrality that must be in place if his agents are to be capable of agreement that does not presuppose shared background convictions. I have argued that Williams, and the cognate account of Charles Taylor, cannot be drawn on to secure such an account. Neither successfully eliminates prevailing background convictions, those that are, in Rawls, associated with comprehensive doctrine, as the source of what a person thinks even in deepest personal reflection. Their approaches contrast with those of John McDowell who views, if I read him correctly, normative thinking as characterised by a higher-level criticality that does not privilege the presuppositions of a particular morality. This degree of criticality is analogous to that which I have argued ought to be associated with the

market participant. Indeed, in conjunction with his thesis on the role of moral education in promoting a neutral take on existing norms, we can present a McDowellian account of the moral education of the young in a market economy. This permits a combination of thoroughgoing commitment to existing institutions, commitment that emanates from comprehensive doctrine, and so engenders stability, while allowing neutral critique in the assessment of political, economic and social institutions.

Conclusion

The neo-Hegelian criticism of Rawls, to the extent that it applies to his conception of the person has been, arguably, refuted by Rawls' reinterpretation of his theory of justice. By saying that his conception of the person is meant to model but not to represent actual evaluative capacities, Rawls effectively side-steps the neo-Hegelian remark. However, the way Rawls constructs impartial assessment has implications for his conception of stability in a well-ordered society. Those who object to existing economic or social policy will, where the dispute cannot be settled with reference to existing interpretations of principles of justice, appeal to the argument made in support of the principles of justice themselves. Because entering into this form of argument requires an imaginative loss of a person's existing identity and circumstances, because the reasoning is "behind the veil", the concerns that motivate the claim must be bracketed, and their occurrent, practical significance removed. Where, then, a person accepts that her concern ought not to be acted upon, where a person accepts that the reasoning that supports the principles of justice will not also support her claim, she must, in order to suppress the practical significance of her concerns, mute, or control the strength of, their normative pull. In order to do this, however, on the conception of reflective criticism Rawls gives us, she must act on herself from outside of her everyday way of thinking, what in Rawls is her comprehensive outlook. In Bernard Williams' well known terminology, she must act on herself in a way that

draws on resources that are outside of her “motivational set”. And because the personal adjustment the person must make does not engage with the bases of her everyday practical orientation, she must root out, or extirpate, those personal concerns that would otherwise inform her action. More generally, a well-ordered society regulated by Rawlsian principles, principles supported in the way he would have us support them, will be characterised by uncritical obedience to existing policies. Stability, in a liberal democratic society whose institutions are sensitive to criticism ought, rather, to be characterised as one in which the potential for dissent is ever-present. There should be, in the way stability is conceptualised, an ongoing tension between persons’ commitment to existing institutions and the desire to see some policies revised or replaced. Rawls fails to capture this tension because, as the neo-Hegelian would agree, his conception of reflective critique requires a detachment that annuls the normative purchase of personal concerns. Unlike the neo-Hegelian criticism, however, this objection cannot be easily rebuffed by Rawls’ appeal to the artificiality of his conception of reflective critique. Rawls wants his picture of stability to be realistic. While the well-ordered society merely models actual democratic societies it, nonetheless, is intended to represent a political settlement that is achievable, “taking men as they are”, in the Rousseauian phrase. That is, the institutions in the well-ordered society, though established by ideal theory, ought to be supportable in a way that takes account of the way real world agents (the way “we”) are likely to respond to institutional restraints. It is not realistic, I suggest, to expect, or, moreover, to entertain the possibility that, actual people will root out concerns that would otherwise lead them to dissent. We, rather, temper, through an effort of will, the

practical pull of personal concerns that, though tempered, retain their status as live concerns.

It is possible to argue, of course, for a conception of the person that accommodates ongoing critical acceptance of institutional rules and policy decisions. Such a conception is available, for example, in the theories of Williams and Taylor, theories that I have considered in the text. However, what Rawls requires, if he is to answer the charge that he misrepresents the problem of social stability, is a conception that allows for a neutrality that does not privilege any particular comprehensive point of view. He must have this if he is to avoid the sort of comprehensive liberalism that he presented in his early work, and a Williams-Taylor type of approach cannot, I have argued, offer this. Drawing on the work of John McDowell, it is possible, however, to work up a reading of Rawls that achieves these two desiderata: a critical capacity that does not privilege comprehensive doctrine; and that retains the vital practical significance of personal concern. This is an appropriate model to associate with the market participant who, I have argued, if he is to be soundly characterised as a market agent, must possess higher-order critical faculties that transcend the divide between neutral and non-neutral, doctrinal thinking. One advantage of such an approach is that it works from within Rawls' own theory. Rather than importing a thesis concerning a conception of critical reflection from outside of his own approach, my method has been to draw on Rawls' presumption of discretionary, market-based exchange. It is possible in this way to leave intact all of Rawls' contentions, and his way of defending them, changing only our interpretation of the reflective capacities of his agents and, relatedly, the nature of the original

position. Rather than taking the latter to be a free-standing model of reflective critique, a reading that links it naturally to the faulty, bipartite conception, we should see the original position as something that is embraced by Rawlsian agents from a point of view that encapsulates both neutral and non-neutral standpoints. If criticality can, as I have argued, be viewed as fundamentally presuppositionless then it would be wrong to suppose that agents must cleanse themselves of their comprehensive doctrines in order to see things impartially. That is not to say that they ought not to imaginatively remove from view their talents, existing status and so on, but this need not lead us to think that they must, thereby, blind themselves to the motive force of their ongoing concerns. This accommodation (of the imaginative adoption of the impartial point of view and the living practical significance of personal concerns) offers, I suggest, a combination of reflective critique and personal embeddedness that promises to mend what is wrong with the representation of social stability in Rawls.

Bibliography

Ake, Claude (1974) "Modernization and Political Instability: A Theoretical Exploration", *World Politics*, Jul., Vol. 26, No. 4, 576-591.

Alchian, Armen A. and Demsetz, Harold (1996) "Production, Information Costs, and Economic Organization", in Peter J. Buckley and Jonathan Michie (eds.) *Firms Organisations and Contracts* (Oxford: Oxford University Press). Originally in *American Economic Review* Vol. 62, 1972, 777-95.

Arrow, Kenneth, J. (1963/1951) *Social Choice and Individual Values* (New York: John Wiley and Sons).

Ahmed, Akbar, S. (1986) No title: Comments in Response to Bêteille (1986), *Current Anthropology*, Vol. 27, No. 2, April, 128.

Alexander, Larry (1985) "Pursuing the Good-Indirectly", *Ethics*, Vol. 95, No. 2, Jan, 315-332.

Anderson, Elizabeth, S. (1999) "What is the Point of Equality", *Ethics*, Vol. 109, Issue 2, Jan., 287-337.

Barry, Brian (1995) "John Rawls and the Search for Stability", *Ethics*, Vol. 105, No. 4, Jul., 874-915.

Beitz, Charles, R. (1983) "Cosmopolitan Ideals and National Sentiment", (in Symposium: International Justice), *The Journal of Philosophy*, Vol. 80, No. 10, 591-600.

_____ (1979) *Political Theory and International Relations* (Princeton: Princeton University Press).

Benhabib, Seyla (1992) *Situating the Self* (London: Routledge).

Bentham, Jeremy (1968/1823) "An Introduction to the Principles of Morals and Legislation" (Chapters 1-4) in Alfred N. Page (ed.) *Utility Theory: A Book of Readings* (London: John Wiley and Sons).

Bergson A. (1968/1938) "A Reformulation of Certain Aspects of Welfare Economics", in Alfred N. Page (ed.) *Utility Theory: A Book of Readings* (London: John Wiley and Sons).

- Berlin, Isaiah** (1976) "Herder and the Enlightenment", in *Vico and Herder: Two Studies in the History of Ideas* (London: The Hogarth Press).
- Béteille, André** (1986) "Individualism and Equality", *Current Anthropology*, Vol. 27, No. 2, April, 121-128.
- Bevir, Mark and O'Brien, David** (2003) "From Idealism To Communitarianism: The Inheritance And Legacy Of John Macmurray", *History Of Political Thought*, Vol. XXIV, No. 2, Summer, 305-329.
- Binmore, K.** (1994) *Playing Fair* (Cambridge, MA.: The MIT Press).
- Bird, G** (1996) "McDowell's Kant: Mind and World", *Philosophy*, Vol. 71, 219-243.
- Blackburn, Simon** (2001) *Ruling Passions: A Theory of Practical Reasoning* (Oxford: Oxford University Press).
- _____ (1985) "Errors and the Phenomenology of Value", in T. Honderich (ed.) *Morality and Objectivity* (Boston: Routledge & Kegan Paul).
- Bohman James** (1992) "The Limits of Rational Choice Explanation" in James S. Coleman and T J Fararo (eds.) *Rational Choice Theory: Advocacy and Critique* (Newbury Park, California: Sage Publications)
- Bordes, George and Le Breton, Michel** (1989) "Arrovian Theorems with Private Alternatives Domains and Selfish Individuals", *Journal of Economic Theory*, 47, 257-281.
- Bordes, George, Le Breton, Michel and Hammond, Peter** (2005) "Social Welfare Functionals on Restricted Domains and in Economic Environments", *Journal of Public Economic Theory*, Feb., Vol. 7, no. 1, 1-25. Previously available in *Stanford University Department of Economics Working Paper*, 1999, No. 97-023.
- Bowie, Andrew** (1996) "The Meaning of the Hermeneutic Tradition in Contemporary Philosophy", in Anthony O'Hear (ed.) *Verstehen and Humane Understanding* (Cambridge: Cambridge University Press).
- Bowles, Samuel** (1998) "Endogenous Preferences: The Cultural Consequences of Markets and Other Economic Institutions", *Journal of Economic Literature*, Vol. 36, No. 1, Mar., 75-111.

- Blackorby, C., Donaldson, D. and Weymark, J. A.** (1984) "Social Choice with Interpersonal Utility Comparisons: A Diagrammatic Introduction", *International Economic Review*, June, 25(2), 325-56.
- Brennan, Timothy J.** (1993) "The Futility of Multiple Utility" *Economics and Philosophy*, 9, 155-164.
- _____ (1989) "A Methodological Assessment of Multiple Utility Frameworks", *Economics and Philosophy*, 5, 189-208
- Broome, John** (1999) "Can a Humean be Moderate?" in *Ethics Out of Economics* (Cambridge: Cambridge University Press).
- _____ (1991) "Utility", *Economics and Philosophy*, 7, 1-12.
- Buchanan, Allen** (1989) "Assessing the Communitarian Critique of Liberalism", *Ethics*, 99, July, 852-882.
- Burnyeat, M.** (1980) "Aristotle on Learning to be Good", in Amelie O. Rorty (ed) *Essays on Aristotle's Ethics* (Berkeley: University of California Press).
- Capitani, Diane N.** (2003) "Ideas of the Good: Iris Murdoch's the Sea, the Sea" *Christianity and Literature*, Vol. 53, 99-108.
- Carens, Joseph** (2000) *Culture, Citizenship, and Community* (Oxford: Oxford University Press).
- Coase, Ronald** (1998) "The New Institutional Economics", *The American Economic Review*, Vol. 88, No. 2, May, 72-74.
- _____ (1937) "The Nature of the Firm", *Economica* 4, 386-405. Reprinted in Peter J. Buckley and Jonathan Michie (eds.) (1996) *Firms Organisations and Contracts* (Oxford: Oxford University Press).
- Cohen, Andrew Jason** (1999) "Communitarianism, 'Social Constitution', and Autonomy", *Pacific Philosophical Quarterly*, 80, 121-135.
- Cohen, Gerald A.** (2003) "Facts and Principles", *Philosophy and Public Affairs*, Summer, Vol. 31, Issue 3, 211-245.
- _____ (2000) *If You're an Egalitarian, How Come You're so Rich?* (Cambridge, Mass.: Harvard University Press). Especially chapters 8, 9, and 10.
- _____ (1997) "Where the Action is: On the Site of Distributive Justice." *Philosophy and Public Affairs*, 26, 3-30.

_____ (1995) "The Pareto Argument for Inequality", *Social Philosophy and Policy*, 12, 160-185.

_____ (1992) "Incentives, Inequality, and Community", in Grethe B. Peterson (ed.) *The Tanner Lectures on Human Values*, Vol. 13 (Salt Lake City: University of Utah Press).

Cohen, Jere (1975) "Moral Freedom Through Understanding in Durkheim", *American Sociological Review*, Vol. 40, No 1, Feb, 104-106.

Coleman J. S. (1986) "Beyond Pareto Optimality" in *Individual Interests and Collective Action* (Cambridge: Cambridge University Press).

Cowell, Frank (2006) *Microeconomics: Principles and Analysis* (Oxford: Oxford University Press).

Cullity, G. (1997) "Introduction", in Garrett Cullity and Berys Gaut (eds.) *Ethics and Practical Reason* (Oxford: Clarendon Press).

D'Aspremont, C. and Gevers, L. (1977) "Equity and the Informational Basis of Collective Choice", *Review of Economic Studies*, 44, Jun., pp. 199-204.

Dahl, Robert and Lindblom, Charles (1992/1953) *Politics, Economics and Welfare* (Brunswick: Transactions Publishers).

Darwall, S. Gibbard, A. and Railton, P. (eds.) (1997) "Introduction", *Moral Discourse and Practice* (Oxford: Oxford University Press).

Davidson, D. (1986) "Empirical Content" in Ernest le Pore (ed) *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson* (Oxford: Blackwell), 320-332.

Dallmayr, Fred, R. (1974) "Toward a Critical Reconstruction of Ethics and Politics", *The Journal of Politics*, Vol. 36, No. 4, Nov., 926-957.

Debreu, Gerard (1972/1959) *Theory of Value: An Axiomatic Analysis of Economic Equilibrium* (Yale: Yale University Press).

Dickenson, Donna (2002) "Black and White and Right All Over?" *University of Birmingham, Centre for Global Ethics*, Inaugural Lecture, 23rd May.

Dilley, Roy (1991) "Conferences: Contesting Markets", *Anthropology Today*, Vol. 7, No. 4, Aug., 14-18.

Drakopoulos, S. A. (1994) "Hierarchical Choice in Economics", *Journal of Economic Surveys*, Vol. 8, No. 2, 133-153.

- Dworkin, Ronald** (2000) *Sovereign Virtue* (Cambridge, Mass.: Harvard University Press).
- (1981a) "What is Equality? Part 1: Equality of Welfare", *Philosophy and Public Affairs*, Vol. 10, No. 3, Summer, 185-246.
- _____ (1981b) "What is Equality? Part 2: Equality of Resources", *Philosophy and Public Affairs*, Vol. 10, No. 4, Autumn, 283-345.
- Elster, J. and Hylland, A.** (eds.) (1986) *Foundations of Social Choice Theory* (Cambridge: Cambridge University Press).
- Elster, J.** (1985a) "Weakness of the Will and the Free Rider Problem", *Economics and Philosophy*, 1, 231-65.
- _____ (1985b) "Introduction", *The Multiple Self*, Jon Elster (ed.) (Cambridge: Cambridge University Press).
- Etzioni, Amitai** (1986) "The Case for a Multiple-Utility Conception", *Economics and Philosophy*, 2, 159-84.
- Flynn, Norman** (2002) *Public Sector Management* (Financial Times/Prentice Hall) 4th edition.
- Foot, Phillipa** (1972) "Morality as a System of Hypothetical Imperatives", *The Philosophical Review*, Jul., Vol. 81, No. 3, 305-316.
- Frank S. Fanselow** (1990) "The Bazaar Economy or How Bizarre is the Bazaar Really?" *Man*, New Series, Vol. 25, No. 2, Jun., 250-265.
- Frankfurt, Harry G.** (1971) "Freedom of the Will and the Concept of a Person", *The Journal of Philosophy*, Vol. LXVIII, No. 1, Jan, 14, 5-20.
- Freeman, Samuel** (2003) "Introduction: John Rawls an Overview", in *The Cambridge Companion to Rawls* (Cambridge: Cambridge University Press).
- _____ (ed.) (1999) *John Rawls: Collected Papers*, (Cambridge Mass.: Harvard University Press).
- Friedman, Milton** (1970) "The Social Responsibility of Business is to Increase its Profits", *The New York Times Magazine*, Sep. 13.
- _____ (1966) "The Methodology of Positive Economics", in *Essays in Positive Economics* (Chicago: University of Chicago Press).
- Gaut, B. and Cullity, G.** (1997) "Introduction", in Gaut and Cullity (eds.) *Ethics and Practical Reason* (Oxford: Clarendon Press).

Gaut, B. (1997) "The Structure of Practical Reason" in Gaut and Cullity (eds.) *Ethics and Practical Reason* (Oxford: Clarendon Press).

Gauthier, David (1987) *Morals By Agreement* (Oxford: Oxford University Press).

Geertz, Clifford (1978) "The Bazaar Economy: Information and Search in Peasant Marketing", *The American Economic Review*, Vol. 68, No. 2, *Papers and Proceedings of the Ninetieth Annual Meeting of the American Economic Association*, May, 28-32.

George, David (1998) "Coping Rationally With Unpreferred Preferences", *Eastern Economic Journal*, Spring, Vol. 24, Issue 2, 181-195.

_____ (1993) "Does the Market Create Preferred Preferences?" *Review of Social Economy*, 51, 323-346.

Giddens, Anthony (1990) "The Consequences of Modernity", (Cambridge: Polity).

Goodin, Robert E. (1988) "What is So Special about Our Fellow Countrymen?", *Ethics*, Vol. 98, No. 4, Jul., 663-686.

_____ (1986) "Laundering Preferences" in Elster and Hylland (eds.) *Foundations of Social Choice Theory* (Cambridge: Cambridge University Press).

Gosepath, Stefan (2003) "The Global Scope of Justice", in Thomas Pogge (ed.) *Global Justice* (Oxford: Blackwell).

Granovetter, Mark (1985) "Economic Action and Social Structure: The Problem of Embeddedness", *The American Journal of Sociology*, Vol. 91, No. 3, Nov., 481-510.

Habermas, Jürgen (1995) "Reconciliation Through the Public Use of Reason: Remarks on John Rawls's Political Liberalism", *The Journal of Philosophy*, Vol. 92, No. 3, Mar., 109-131.

_____ (1990) *Moral Consciousness and Communicative Action* (Cambridge: Polity).

Hahnel, R. and Albert, M. (1990) *Quiet Revolution in Welfare Economics*, (Princeton: Princeton University Press).

Haldane, John (1996) "Rational and Other Animals", in Anthony O'Hear (ed.) *Verstehen and Humane Understanding* (Cambridge: Cambridge University Press).

- Harcourt, Edward** (1998) "Integrity, Practical Deliberation and Utilitarianism", *The Philosophical Quarterly*, Vol. 48, No 191, 189-198.
- Hare, Richard M.** (1997) *Sorting Out Ethics* (Oxford: Clarendon Press).
- _____ (1981) *Moral Thinking* (Oxford: Oxford University Press).
- _____ (1962) *Freedom and Reason* (Oxford: Oxford University Press).
- _____ (1952) *The Language of Morals* (Oxford: Clarendon).
- Hargreaves Heap, S., Hollis, M., Lyons, B., Sugden, R., Weale, A.** (1992) *The Theory of Choice: A Critical Guide* (London: Blackwell).
- Hahn, Frank** (ed.) (1979) *Philosophy and Economic Theory* (Oxford: Oxford University Press).
- Harsanyi J C** (1982) "Morality and the Theory of Rational Behaviour" in A. Sen and B. Williams (eds.) *Utilitarianism and Beyond* (Cambridge, England: Cambridge University Press)
- _____ (1976a) "Can the Maximin Principle Serve as a Basis for Morality?: A Critique of John Rawls' Theory", in *Essays on Ethics, Social Behaviour and Scientific Explanation* (Dordrecht: Kluwer).
- _____ (1976b) "Cardinal Utility in Welfare Economics and in the Theory of Risk Taking", in *Essays on Ethics, Social Behaviour and Scientific Explanation* (Dordrecht: Kluwer).
- _____ (1976c) "Cardinal Welfare, Individualist Ethics and Interpersonal Comparisons of Utility", in *Essays on Ethics, Social Behaviour and Scientific Explanation* (Dordrecht: Kluwer).
- Hausheer, Roger** (1996) "Three Major Originators of the Concept of *Verstehen*: Vico, Herder, and Schleiermacher", in Anthony O'Hear (ed.) *Verstehen and Humane Understanding* (Cambridge: Cambridge University Press).
- Hausman, Daniel and McPherson Michael** (1996) *Economic Analysis and Moral Philosophy* (Cambridge: Cambridge University Press).
- Hekman, Susan** (1992) "The Embodiment of the Subject: Feminism and the Communitarian Critique of Liberalism", *The Journal of Politics*, Vol. 54, No. 4, Nov., 1098-1119.

Hempel Carl (1965) "The Concept of Rationality and the Logic of Explanation by Reasons", in *Aspects of Scientific Explanation* (New York: The Free Press), 463-486.

Henry M. Oliver, Jr. (1946) "Pure Competition and Coercion", *Ethics*, Vol. 57, No. 1, Oct., 65-66.

_____ (1955) "Attitudes Toward Market and Political Self-Interest", *Ethics*, Vol. 65, No. 3, Apr., 171-180.

Hirschman, Albert O. Hirschman (1992) *Rival Views of Market Society* (Cambridge Mass.: Harvard University Press).

_____ (1985) "Against Parsimony: Three Easy Ways of Complicating Some Categories of Economic Discourse", *Economics and Philosophy*, 1, No. 1. Reprinted in Hirschman (1992).

Hobbes, Thomas (1651/1985) *Leviathan* (London: Penguin Books).

Hollis, Martin (1995) "The Shape of a Life", in J. E. J. Altham and Ross Harrison (eds.) *World, Mind, and Ethics: Essays in the Ethical Philosophy of Bernard Williams* (Cambridge: Cambridge University Press).

_____ (1994) "Autonomy", in *The Theory of Choice*, Shaun Hargreaves Heap, Martin Hollis, Bruce Lyons, Robert Sugden and Albert Weale (Oxford: Blackwell, second edition) (esp. 80-85).

_____ (1987) *The Cunning of Reason* (Cambridge, England: Cambridge University Press) Chs. 2 & 6.

How, Alan (1990) *The Habermas-Gadamer Debate and the Nature of the Social* (Aldershot: Avebury).

Hoy, David (1975) "Literary History: Paradigm or Paradox", *Yale French Studies*, No. 52, *Graphesis: Perspectives in Literature and Philosophy*, 268-286.

Hume, David (1983/1751) *An Enquiry Concerning the Principles of Morals* (Indianapolis: Hackett Publishing Company).

_____ (1978/1740) *A Treatise of Human Nature* (Oxford: Clarendon Press).

Hurwicz L., Schmeidler D., and Sonnenschein H. (1985) *Social goals and social organisation: essays in memory of Elisha Pazner* (Cambridge University Press).

Hylland A (1986) "The Purpose and Significance of Social Choice Theory: Some General Remarks and an Application to the 'Lady Chatterley problem'", in Elster, J. and Hylland, A. *Foundations of Social Choice Theory* (Cambridge: Cambridge University Press).

Kant, Immanuel (1997/1788) *Critique of Practical Reason* (Cambridge: Cambridge University Press).

_____ (1998/1785) *Groundwork for the Metaphysics of Morals* (Cambridge: Cambridge University Press).

_____ (1998/1781) *Critique of Pure Reason* (Cambridge: Cambridge University Press).

Kay, John (2004) *The Truth About Markets* (London: Penguin Books).

_____ (2003) "The Embedded Market", in Anthony Giddens (ed.) *The Progressive Manifesto* (London: Polity).

Kelly, Erin (2000) "Habermas on Moral Justification", *Social Theory and Practice*, Summer, Vol. 26, 223-249.

Kelly, Michael (1990) "MacIntyre, Habermas, and Philosophical Ethics", in Michael Kelly (ed.) *Hermeneutics and Critical Theory in Ethics and Politics*, (Cambridge, Mass.: MIT Press).

Korsgaard, Christine (1998) "Motivation, Metaphysics and the Value of the Self: A Reply to Ginsborg, Guyer and Schneewind", *Ethics*, Vol. 109, No. 1, Oct, 49-66.

_____ (1986) "Skepticism About Practical Reason", *The Journal of Philosophy*, Vol. 83, No. 1, Jan., 5-25.

Kymlicka, Will (1991) *Liberalism, Community, and Culture* (Oxford: Clarendon Press).

_____ (1990) *Contemporary Political Philosophy*, (Oxford: Clarendon Press). (Second edition: 2002).

Lammi, Walter (1991) "Hans-Georg Gadamer's "Correction" of Heidegger", *Journal of the History of Ideas*, Vol. 52, No. 3, Jul., 487-507.

Larmore, Charles (2002) "Attending to Reasons", in Nicholas Smith (ed.) *Reading McDowell: On Mind and World* (London: Routledge).

_____ (1999) "The Moral Basis of Political Liberalism", *The Journal of Philosophy*, Vol. 96, No. 12, Dec., 599-625.

_____ (1996) "The Foundations of Modern Democracy: Reflections on Jürgen Habermas", in *The Morals of Modernity* (Cambridge: Cambridge University Press).

_____ (1992/1987) *Patterns of Moral Complexity* (New York: Cambridge University Press).

Lawrence, Gavin (1995) "The Rationality of Morality", in, Rosalind Hursthouse, Gavin Lawrence, Warren Quinn (eds.) *Virtues and Reasons: Philippa Foot and Moral Theory* (Oxford: Clarendon Press).

Le Pore, Ernest and McLaughlin, Brian (1985) "Actions, Reasons, Causes, and Intentions" in Le Pore and McLaughlin (eds.) *Actions and Events: Perspectives on the Philosophy of Donald Davidson* (Oxford: Basil Blackwell).

Lindblom, Charles (2002) *The Market System*, (New Haven & London: Yale University Press/Yale Nota Bene).

Lutz, Mark A. (1993) "The Utility of Multiple Utility: A Comment on Brennan" *Economics and Philosophy*, 9, 145-154.

MacIntyre, Alasdair (1984) "Is Patriotism a Virtue?", *The Lindley Lecture* (Lawrence: University of Kansas, Department of Philosophy). Reprinted in Richard J. Arneson (ed.) *Liberalism, Volume III* (Aldershot: Elgar, 1992), 246-263.

_____ (1987/1981) *After Virtue* (London: Duckworth). Second edition. Part of this, "The Virtues, the Unity of a Human Life, and the Concept of a Tradition", is excerpted in Michael Sandel (ed.) *Liberalism and its Critics* (Oxford: Basil Blackwell).

_____ (1983) "Moral Arguments and Social Context", *The Journal of Philosophy*, Vol. 80, No. 10, Oct., 590-591.

Mackie, J. L. (1990/1977) *Ethics: Inventing Right and Wrong* (London: Penguin). Second edition.

Macleod, Colin, M. (1998) *Liberalism, Justice, and Markets: A Critique of Liberal Equality* (Oxford: Oxford University Press).

Mara, Gerald. M. (1985) "After Virtue, Autonomy: Jürgen Habermas and Greek Political Theory", *The Journal of Politics*, Vol. 47, No. 4, Nov., 1036-1061.

- Mason, Andrew** (1997) "Special Obligations to Compatriots", *Ethics*, Vol. 107, No. 3, Apr., 427-447.
- McCarthy, Thomas** (1994) "Kantian Constructivism and Reconstructivism: Rawls and Habermas in Dialogue", *Ethics*, Vol. 105, No. 1, Oct., 44-63.
- _____ (1980) "Rationality and Relativism in Habermas' Critical Theory", Vol. 14, No. 1, *A. P. A. Western Division Meetings*, Mar., 175-176.
- McPherson, Michael S.** (1982) "Mill's Moral Theory and the Problem of Preference Change" *Ethics*, Vol. 92, No. 2, 252-273.
- McDowell, John** (1998a) "Lecture III: Intentionality as a Relation", *The Journal of Philosophy*, Vol. 95, No. 9, Sep., 471-491.
- _____ (1998b) "Lecture II: The Logical Form of an Intuition", *The Journal of Philosophy*, Vol. 95, No. 9, Sep., 451-470.
- _____ (1998c) "Lecture I: Sellars on Perceptual Experience", *The Journal of Philosophy*, Vol. 95, No. 9, Sep., 431-450.
- _____ (1998d) "One Strand in the Private Language Argument", in *Mind, Value and Reality* (Cambridge, Mass.: Harvard University Press).
- _____ (1995) "Might There be External Reasons?", in J. E. J. Althom and Ross Harrison (eds.) *World, Mind, and Ethics: Essays on the Ethical Philosophy of Bernard Williams* (Cambridge: Cambridge University Press).
- _____ (1995) "Two Sorts of Naturalism", in R. Hursthouse, G Lawrence and W Quinn (eds.) *Virtues and Reasons* (Oxford: Oxford University Press).
- _____ (1994) *Mind and World* (Harvard University Press).
- _____ (1994) "The Content of Perceptual Experience", *The Philosophical Quarterly*, Vol. 44, No. 175, Apr., pp. 190-205.
- _____ (1985) "Values and Secondary Qualities", in Ted Honderich, ed., *Morality and Objectivity* (Boston: Routledge and Kegan Paul).
- _____ (1981) "Non-Cognitivism and Rule-Following", in Steven H. Holtzman and Christopher Leach (eds.) *Wittgenstein: to Follow a Rule* (London, Boston and Henley: Routledge & Kegan Paul).
- _____ (1979) "Virtue and Reason", *The Monist*, Vol 62, No. 3, July, 331-350.

- Nietzsche, F. (1978) "Are Moral Requirements Hypothetical Imperatives?" *Proceedings of the Aristotelian Society Supplementary Volume*, 13-29.
- McDowell, John, et al** (1998) "Book Symposium: Mind and World", *Philosophy and Phenomenological Research*, Vol. 58, No. 2, Jun., 365-486.
- Mill, John Stuart** (1972/1863) *Utilitarianism*, in H. B. Acton (ed.) *Utilitarianism, Liberty, Representative Government* (London: J M Dent and Sons).
- O'Neill, On (1972/1859) *On Liberty*, in H. B. Acton (ed.) *Utilitarianism, Liberty, Considerations on Representative Government* (London: J M Dent and Sons).
- Bleeden, V. (1965/1844) "On the Definition of Political Economy: and on the Method of Investigation Proper to it", Essay V in *Essays on Some Unsettled Questions of Political Economy* (1844) cited in the Introduction to *Collected Works of John Stuart Mill: Introduction to the Principles of Political Economy, Books I and II*, V. Bleeden and J. Robson (eds.) 1965 (University of Toronto Press: Routledge and Kegan Paul).
- Miller, David** (1990) *Market, State and Community: Theoretical Foundations of Market Socialism* (Oxford: Oxford University Press).
- Parfit, D. (1988) "The Ethical Significance of Nationality", *Ethics*, Vol. 98, No. 4, Jul., 647-662.
- Mortimore, G. W.** (1976) "Rational Action", in *Rationality and the Social Sciences*, S. I. Benn and G. W. Mortimore (eds.) (London: Routledge and Kegan Paul).
- Murdoch, Iris** (1970) *The Sovereignty of Good* (London: Routledge).
- Murphy, Liam** (1998) "Institutions and the Demands of Justice", *Philosophy and Public Affairs*, 27, 251-291.
- Nagel, Thomas** (1970) *The Possibility of Altruism* (Princeton: Princeton University Press)
- Neurath, Otto** (1959) "Protocol Sentences", in A. J. Ayer (ed.) *Logical Positivism* (New York: Free Press).
- Nida-Rumelin, Julian** (1993) "Practical Reason, Collective Rationality and Contractarianism", in D. Gauthier and R. Sugden (eds.) *Rationality, Justice and the Social Contract* (London: Harvester Wheatsheaf).

Nielsen, Kai (1983) "Global Justice and the Imperatives of Capitalism" (in Symposium: International Justice), *The Journal of Philosophy*, Vol. 80, No. 10, 608-610.

Nozick, R. (1974) *Anarchy, State and Utopia* (New York: Basic Books)

Oakley, J. (2001) "Humane After All?" *Times Literary Supplement*, July 6, 10-11.

O'Neill, Onora (2002) "Global Justice: Whose Obligations?" *Centre for the Study of Global Ethics*, University of Birmingham, 23rd May 2002, available at <http://www.globalethics.bham.ac.uk/Activities/ONeill.htm> (April 2004)

_____ (1988) "Ethical Reasoning and Ideological Pluralism", *Ethics*, 98, July, 705-722.

Page, Alfred N. (ed.) (1968) *Utility Theory: A Book of Readings* (London: John Wiley and Sons).

Parsons, Talcott (1975) "Comment on 'Parsons' Interpretation of Durkheim" and on "Moral Freedom Through Understanding in Durkheim", *American Sociological Review*, Vol. 40, No. 1, Feb., 106-111.

Payutto, P. A. (1994) *Buddhist Economics* (Bangkok: Buddhadhamma Foundation); second edition.

Pogge, Thomas (2003) *Global Justice* (Oxford: Blackwell).

(2002a) "Moral Universalism and Global Economic Justice", in Thomas Pogge *World Poverty and Human Rights* (Cambridge: Polity).

_____ (2002b) "The Bounds of Nationalism", in Thomas Pogge *World Poverty and Human Rights* (Cambridge: Polity).

_____ (2002c) "Cosmopolitanism and Sovereignty", in Thomas Pogge *World Poverty and Human Rights* (Cambridge: Polity).

_____ (2000) "On the Site of Distributive Justice", *Philosophy and Public Affairs*, 29, No. 2, 137-169.

Polanyi, Karl (2001/1944) *The Great Transformation: The Political and Economic Origins of Our Time* (Boston: Beacon Press).

Pope, Whitney (1975) "Parsons on Durkheim Revisited", *American Sociological Review*, Vol. 40, No. 1, Feb., 111-115.

_____ (1973) "Classic on Classic: Parson's Interpretation of Durkheim", *American Sociological Review*, Vol. 38, Aug., 399-415.

Rawls, John (2001) *Justice as Fairness: A Restatement* (Cambridge Mass.: The Belknap Press of Harvard University Press).

_____ (2001) *The Law of Peoples* (Cambridge, Mass.: Harvard University Press), third printing.

_____ (1997) "The Idea of Public Reason Revisited", *University of Chicago Law Review*, 64, Summer, No. 3, 765-807. Reprinted in *The Law of Peoples* (Cambridge, Mass.: Harvard University Press), 129-180.

_____ (1996/1993) *Political Liberalism* (New York: Columbia University Press). Second edition.

_____ (1995) "Political Liberalism: Reply to Habermas", *The Journal of Philosophy*, Vol. 92, No. 3, Mar., pp. 132-180.

_____ (1989) "Themes in Kant's Moral Philosophy", in *Kant's Transcendental Deductions*, Eckart Forster (ed.) (Stanford: Stanford University Press), 81-113.

_____ (1985) "Justice as Fairness: Political Not Metaphysical", *Philosophy and Public Affairs*, Vol. 14, No. 3, Summer, 223-251. Reprinted in Freeman (ed.) (1999).

_____ (1999/1982) "Social Unity and Primary Goods" in Freeman (ed.) (1999), reprinted from A. Sen and B. Williams (eds.) (1982) *Utilitarianism and Beyond* (Cambridge: Cambridge University Press).

_____ (1980) "Kantian Constructivism in Moral Theory", *The Journal of Philosophy* 77, No 9 Sep., pp515-572, and selected sections in "Moral Discourse and Practice" (1997) S. Darwall, A. Gibbard, and P. Railton (eds.) (Oxford: Oxford University Press), 247-2.

_____ (1973/1971) *A Theory of Justice* (Oxford: Oxford University Press).

_____ (1999/1971) *A Theory of Justice* (Oxford: Oxford University Press).

_____ (2000) *Lectures on the History of Moral Philosophy* (Cambridge, MA: Harvard University Press).

_____ (1957) "Justice as Fairness", *The Journal of Philosophy*, Vol. 54, No. 22, Oct. American Philosophical Association Eastern Division: Symposium

Papers to be Presented at the Fifty-Fourth Annual Meeting, Harvard University, December 27-29, 1957. (Oct. 24, 1957), 653-662.

Raz, Joseph (1999) *Engaging Reason* (Oxford: Oxford University Press).

Reiman, Jeffrey (1987) "Exploitation, Force, and the Moral Assessment of Capitalism: Thoughts on Roemer and Cohen", *Philosophy and Public Affairs*, Winter, Vol. 16, No. 1, 3-41.

Resnick, Michael (1986) "Impartial Welfarism and the Concept of a Person", *Erkenntnis* 25: 47-60.

Ridge, Michael (1998) "Hobbesian Public Reason" *Ethics*, 108, April, 538-568.

Riley, J. (1988) *Liberal Utilitarianism: Social Choice Theory and J S Mill's Philosophy* (Cambridge: Cambridge University Press).

Rizvi, Turab (2001) "Preference Formation and the Axioms of Choice", *Review of Political Economy*, April, Issue 2, Vo. 13, 141-159.

Rorty, Richard (1980) *Philosophy and the Mirror of Nature* (Princeton: Princeton University Press).

Ross, J. J. (1970) *The Appeal to the Given* (London: George Allen and Unwin Ltd).

Rousseau, J. J. (1950/1762) *The Social Contract* (London: J. M. Dent).

Russell, Bertrand (1945) *A History of Western Philosophy* (New York: Simon and Schuster).

Sandel, Michael (1996) *Democracy's Discontent* (Cambridge, Mass.: The Belknap Press of Harvard University Press).

_____(1998/1982) *Liberalism and the Limits of Justice* (Cambridge: Cambridge University Press). Second edition.

Scanlon, T. M. (1982) "Contractarianism and Utilitarianism" in A. Sen and B. Williams (eds.) *Utilitarianism and Beyond* (eds.) (Cambridge: Cambridge University Press).

Scarre, Geoffrey (1996) *Utilitarianism* (London: Routledge).

Scheffler, Samuel (2003) "What is Egalitarianism", *Philosophy and Public Affairs*, Winter, Vol. 31, Issue 1, 5-39.

Scheffler, Samuel (1988a) *Consequentialism and its Critics* (Oxford: Oxford University Press).

Scheffler, Samuel (1988b) "Agent-Centred Restrictions, Rationality, and the Virtues", in Samuel Scheffler (ed.) *Consequentialism and its Critics* (Oxford: Oxford University Press).

_____(1994/1982) *The Rejection of Consequentialism* (Oxford: Oxford University Press).

Schelling T. (1984) "The Intimate Contest for Self Command", in *Choice and Consequence* (Cambridge, Mass.: Harvard University Press).

Scholte, Jan Aart (2000) *Globalization: A Critical Introduction* (London: Palgrave Macmillan).

Seanor, D. and Fotion, N. (1988) *Hare and Critics* (Oxford: Clarendon)

Sellars, Wilfrid (1997/1956) *Empiricism and the Philosophy of Mind* (Cambridge Mass.: Harvard University Press).

Sen, Amartya (2000) *Development as Freedom* (New York: Anchor Books).

_____(1999a) "The Possibility of Social Choice", *The American Economic Review*, June, 349-378.

_____(1999b) *Reason Before Identity* (Oxford: Oxford University Press).

_____(1997) "Maximisation and the Act of Choice", *Econometrica*, Vol. 65, No. 4, July, 745-779

_____(1995) "Is the Idea of Purely Internal Consistency of Choice Bizarre?" in J. E. J. Altham & Ross Harrison (eds.) *World, Mind and Ethics* (Cambridge: Cambridge University Press).

_____(1986) "Social Choice Theory" in K. Arrow and M. Intriligator (eds.) *Handbook of Mathematical Economics*, Vol III (Amsterdam: North-Holland).

_____(1982a) "On weights and Measures: Informational Constraints in Social Welfare Analysis", in *Choice Welfare and Measurement* (Oxford: Basil Blackwell).

_____(1982b) *Choice Welfare and Measurement* (Oxford: Basil Blackwell).

_____(1982c) "Personal Utilities and Public Judgements: Or What's Wrong With Welfare Economics?" in *Choice Welfare and Measurement* (Oxford: Basil Blackwell).

_____(1980) "Plural Utility", *Proceedings of the Aristotelian Society*, New Series, Vol Lxxxi, 193-215.

_____(1979) "Utilitarianism and Welfarism" *The Journal of Philosophy*, 9, Sep., 463-489.

_____(1977) "Rational Fools: A Critique of the Behavioural Foundations of Economic Theory", *Philosophy and Public Affairs*, 6 Summer 317-44 and in *Choice, Welfare and Measurement* (1982) (Cambridge, Mass.: MIT Press).

_____(1974) "Choice, Orderings and Morality", in S. Korner (ed.) *Practical Reason* (Oxford: Blackwell), and in "Choice, Welfare and Measurement", A. Sen (1982) (Cambridge, Mass.: MIT Press).

_____(1973) "Behaviour and the Concept of Preference", *Economica*, 40, August, 241-59 and in *Choice Welfare and Measurement*, A. Sen (1982) (Cambridge, Mass.: MIT Press).

_____(1970a) *Collective Choice and Social Welfare* (San Francisco: Holden Day).

_____(1970b) "The Impossibility of a Paretian Liberal" *Journal of Political Economy*, 78, 152-157.

Shue, Henry (1983) "The Burdens of Justice" (in Symposium: International Justice) *The Journal of Philosophy*, Vol. 80, No. 10, 600-608.

Sidgwick, Henry (1901) *The Methods of Ethics* (London: Macmillan & Co. Ltd.).

Singer, P. (2002) *One World: The Ethics of Globalization* (New Haven: Yale University Press).

_____(1973) "The Triviality of the Debate over "Is-Ought" and the Definition of "Moral"", *American philosophical Quarterly*, 10.

_____(1972) "Famine, Affluence, and Morality", *Philosophy and Public Affairs*, Vol. 1, No. 3, Spring, 229-243.

Smith, Michael (1994) *The Moral Problem* (London: Blackwell).

Smith, Nicholas H. (2002) *Reading McDowell: On Mind and World* (London: Routledge).

Soffer, Gail (2003) "Revisiting the Myth: Husserl and Sellars on the Given", *The Review of Metaphysics*, Vol. 57, 301-337.

Steedman, Ian and Krause, Ulrich (1985) "Goethe's Faust, Arrow's Possibility Theorem and the Individual Decision Taker", in Jon Elster (ed.) *The Multiple Self* (Cambridge, England: Cambridge University Press)

- Stocker, Michael** (1976) "The Scizophrenia of Modern Ethical Theories", *The Journal of Philosophy*, Vol. 73, Issue 14, On Motives and Morals, Aug, 453-466.
- Strawson, Peter** (1962) "Freedom and Resentment", *Proceedings of the British Academy*, Vol. XLVIII. Reprinted in *Freedom and Resentment and Other Essays* (1974) (London: Methuen).
- available at <http://www.ucl.ac.uk/~uctytho/dfwstrawson1.htm> (Jan 2005).
- Sunstein, Cass** (1997a) "Preferences and Politics", in *Free Markets and Social Justice* (Oxford: Oxford University Press).
- _____ (1997b) "Social Norms and Social Roles", in *Free Markets and Social Justice* (Oxford: Oxford University Press).
- Swedberg, Richard** (2003) *Principles of Economic Sociology* (Princeton: Princeton University Press).
- Taylor, Charles** (2002) "Foundationalism and the Inner-Outer Distinction" in Nicholas H. Smith (ed.) *Reading McDowell* (London: Routledge).
- _____ (1989) *Sources of the Self: The Making of Modern Identity* (Cambridge Mass.: Harvard University Press).
- _____ (1985a) "The Diversity of Goods", in *Philosophy and the Human Sciences* (Cambridge: Cambridge University Press).
- _____ (1985b) "What is Human Agency" in *Human Agency and Language: Philosophical Papers 1* (Cambridge: Cambridge University Press). Originally in T. Mischel (ed.) (1977) *The Self* (Oxford: Blackwell).
- _____ (1985c) "Self-interpreting Animals", in *Human Agency and Language: Philosophical Papers 1* (Cambridge: Cambridge University Press) (originally 1977, unpublished).
- _____ (1985d) "Rationality" in *Philosophy and the Human Sciences: Philosophical Papers 2* (Cambridge: Cambridge University Press).
- _____ (1979) *Hegel and Modern Society* (Cambridge: Cambridge University Press).
- Thaler, Richard and Shefrin, H. M.** (1981) "The Economic Theory of Self-Control", *Journal of Political Economy*, vol. 89, No. 2, 392-406.
- Tungoddden, Bertil** (2000) "Egalitarianism: Is Leximin the Only Option?" *Economics and Philosophy*, 16, 229-245.

Vanberg Viktor, J. (1994) "Rules and Choice in Economics and Sociology", in *Rules and Choice in Economics*, (London: Routledge).

Varian, Hal, R. (1992) *Microeconomic Analysis* (New York: W. W. Norton). Third edition.

Walzer, Michael (2004) *Politics and Passion: Toward a More Egalitarian Liberalism* (London and New Haven: Yale University Press).

_____ (1994) *Thick and Thin: Moral Argument at Home and Abroad* (Notre Dame: University of Notre Dame Press).

_____ (1983) *Spheres of Justice: A Defence of Pluralism and Equality* (Oxford: Martin Robertson)

Warnke, Georgia (2002) "Hermeneutics, Ethics, and Politics", in Robert J. Dorstal, *The Cambridge Companion to Gadamer* (Cambridge: Cambridge University Press).

Watkins, J. W. N. (1974) "Comment: Self-interest and Morality," in Stephen Korner (ed.) *Practical Reason* (Oxford: Basil Blackwell).

Wenar, Leif (2003) "Contractualism and Global Economic Justice," in Thomas Pogge (ed.) *Global Justice* (Oxford: Blackwell).

Weymark, John, A. (1998) "Welfarism on Economic Domains", *Mathematical Social Sciences*, 36 , 251-268. Also available as Department of Economics, The University of Columbia, Discussion Paper, No. 97-11. Available at <http://web.arts.ubc.ca/econ/> (Jun. 2003).

White, Stephen, K. (1980) "Reason and Authority in Habermas: A Critique of the Critics", *The American Political Science Review*, Vol. 74, No. 4, Dec., 1007-1017.

Wiggins, David (1998/1987) *Needs, Values, Truth* (Oxford: Basil Blackwell)

Williams, Andrew (1998) "Incentives, Inequality and Publicity", *Philosophy and Public Affairs*, 27, No 3, 225-247.

Williams, Bernard (1995) "Replies", in J. E. J. Altham & Ross Harrison (eds.) *World, Mind, and Ethics*, 185-224 (Cambridge: Cambridge University Press).

_____ (1985) *Ethics and the Limits of Philosophy* (London: Fontana Press/Collins).

_____ (1981a) "Persons, Character and Morality", in *Moral Luck* (Cambridge: Cambridge University Press).

_____ (1981b) "Internal and External Reasons", in *Moral Luck* (Cambridge: Cambridge University Press). Originally in Ross Harrison (ed.) (1980) *Rational Action* (Cambridge: Cambridge University Press).

_____ (1981c) "Utilitarianism and Moral Self-indulgence", in *Moral Luck* (Cambridge: Cambridge University Press).

_____ (1973) "A Critique of Utilitarianism", in J. J. C. Smart and Bernard Williams *Utilitarianism: For and Against* (Cambridge: Cambridge University Press).

Williamson, Oliver E. (2002) "The Theory of the Firm as Governance Structure: From Choice to Contract", *The Journal of Economic Perspectives*, Aug., Volume 16, Number 3, 171-195.

_____ (2000) "The New Institutional Economics: Taking Stock, Looking Ahead", *Journal of Economic Literature*, Vol. XXXVIII, Sept., 595-613.

_____ (1998) "The Institutions of Governance", *The American Economic Review*, Vol. 88, No. 2, Papers and Proceedings of the Hundred and Tenth Annual Meeting of the American Economic Association, May, 75-79.

_____ (1996) "Transaction-Cost Economics: The Governance of Contractual Relations", in Peter J. Buckley and Jonathan Michie (eds.) *Firms Organisations and Contracts* (Oxford: Oxford University Press). Originally in *Journal of Law and Economics*, 1976, XXII, No. 2, 233-261.

_____ (1994) "Transaction Cost Economics and Organization Theory", in Neil J. Smelser and Richard Swedberg (eds.) *The Handbook of Economic Sociology* (Princeton: Princeton University Press).

_____ (1975) *Markets and Hierarchies: Analysis and Antitrust Implications* (London: The Free Press).

Wingebach, Ed (1999) "Unjust Context: The Priority of Stability in Rawls's Contextualized Theory of Justice", *American Journal of Political Science*, Jan., Vol. 43, No. 1, 213-232.

Winship, Christopher and Rosen, Sherwin (1988) "Introduction: Sociological and Economic Approaches to the Analysis of Social Structure", *The American*

Journal of Sociology, Vol. 94, Supplement: Organizations and Institutions: Sociological and Economic Approaches to the Analysis of Social Structure, S1-S16.

Wolff, Jonathan (1991) *Robert Nozick: Property, Justice, and the Minimal State* (Stanford: Stanford University Press).

_____ (1998) "Fairness, Respect, and the Egalitarian Ethos", *Philosophy and Public Affairs*, Vol. 27, No. 2, Spring, 97-122.

Wolff, Robert Paul (1977) *Understanding Rawls* (Princeton: Princeton University).

Wolf, S. (1982) "Moral Saints," *Journal of Philosophy*, 79, 413-39.

Wrong, Dennis (1961) "The Oversocialized Conception of Man in Modern Sociology", *American Sociological Review*, Vol. 26, No. 2, Apr., 183-193.

LC

REFERENCE ONLY



UNIVERSITY OF LONDON THESIS

Degree phd

Year 2007

Name of Author CRAIG ANDREW

DUCKWORTH

COPYRIGHT

This is a thesis accepted for a Higher Degree of the University of London. It is an unpublished typescript and the copyright is held by the author. All persons consulting the thesis must read and abide by the Copyright Declaration below.

COPYRIGHT DECLARATION

I recognise that the copyright of the above-described thesis rests with the author and that no quotation from it or information derived from it may be published without the prior written consent of the author.

LOAN

Theses may not be lent to individuals, but the University Library may lend a copy to approved libraries within the United Kingdom, for consultation solely on the premises of those libraries. Application should be made to: The Theses Section, University of London Library, Senate House, Malet Street, London WC1E 7HU.

REPRODUCTION

University of London theses may not be reproduced without explicit written permission from the University of London Library. Enquiries should be addressed to the Theses Section of the Library. Regulations concerning reproduction vary according to the date of acceptance of the thesis and are listed below as guidelines.

- A. Before 1962. Permission granted only upon the prior written consent of the author. (The University Library will provide addresses where possible).
- B. 1962 - 1974. In many cases the author has agreed to permit copying upon completion of a Copyright Declaration.
- C. 1975 - 1988. Most theses may be copied upon completion of a Copyright Declaration.
- D. 1989 onwards. Most theses may be copied.

This thesis comes within category D.

☐

This copy has been deposited in the Library of

UCL

☐

This copy has been deposited in the University of London Library, Senate House, Malet Street, London WC1E 7HU.